THE AUTOMATIC AND QUANTITATIVE ANALYSIS OF INTERFEROMETRIC AND OPTICAL FRINGE DATA

by

David Peter Towers

Submitted for the Degree of
Doctor of Philosophy
in Engineering
at Warwick University

Originally Submitted
7th March 1991,
Revised 28th April 1991

Department of Engineering
Warwick University
Coventry
England

Abstract

Optical interference techniques are used for a wide variety of industrial measurements. Using holographic interferometry or electronic speckle pattern interferometry, whole field measurements can be made on diffusely reflecting surfaces to sub-wavelength accuracy.

Interference fringes are formed by comparing two states of an object. The interference phase contains information regarding the optical path difference between the two object states, and is related to the object deformation. The automatic extraction of the phase is critical for optical fringe methods to be applied as a routine tool. The solution to this problem is the main topic of the thesis.

All stages in the analysis have been considered : fringe field recording methods, reconstructing the data into a digital form, and automatic image processing algorithms to solve for the interference phase.

A new method for reconstructing holographic fringe data has been explored. This produced a system with considerably reduced sensitivity to environmental changes. An analysis of the reconstructed fringe pattern showed that most errors in the phase measurements are linear. Two methods for error compensation are proposed. The optimum resolution which can be attained using the method is *lamda*/90, or 4 nanometers.

The fringe data was digitised using a framestore and solid state CCD camera. The image processing followed three distinct stages : filtering the input data, forming a 'wrapped' phase map by either the quasi-heterodyne analysis or Fourier transform method, and phase unwrapping. The primary objective was to form a fully automatic fringe analysis package, applicable to general fringe data.

Automatic processing has been achieved by making local measurements of fringe field characteristics. The number of iterations of an averaging filter is optimised according to a measure of the fringe's signal to noise. In phase unwrapping it has been identified that discontinuities in the data are more likely in regions of high spatial frequency fringes. This factor has been incorporated into a new algorithm where regions of discontinuous data are isolated according to local variations in the fringe period and data consistency.

These methods have been found to give near optimum results in many cases. The analysis is fully automated, and can be performed in a relatively short time, $\approx$ 10 minutes on a SUN 4 processor.

Applications to static deflections, vibrating objects, axisymmetric flames and transonic air flows are presented. Static deflection data from both holographic interferometry and ESPI is shown.

The range of fringe fields which can be analysed is limited by the resolution of the digital image data which can be obtained from commercially available devices. For the quantitative analysis of three dimensional flows, the imaging of the fringe data is difficult due to large variations in localisation depth. Two approaches to overcome this problem are discussed for the specific case of burner flame analysis.

Contents

List of Figures

Chapter 1

Introduction

Interferometric methods were established over 100 years ago with the work of Newton, Young, Michelson, Mach, and Zehnder. Standard interferometric methods were limited to examining specularly reflecting surfaces. This restriction was removed approximately 25 years ago by the discovery of holographic interferometry by Powell and Stetson [1]. Since then many developments have occurred in the techniques used to record and interpret the information contained in an interferogram.

Concurrently the range of applications for which the methods offer a solution has grown. These applications include :

i) non-destructive testing (NDT) [2, 3],

ii) strain analysis due to mechanical or thermal loading [4, 5, 6],

iii) modal analysis of vibrations [7, 8, 9],

iv) profile measurement [10, 11, 12],

v) phase object analysis including compressible flows [13, 7, 14], temperature measurement [15, 16], and plasma diagnostics [17].

In every case, the required measurement parameter is encoded as a phase change of the illumination. When two (or more) light waves are superimposed a fringe pattern (or interferogram) is produced.

The results of an interferometric experiment are often required by non-optical specialists. Hence there is a need for methods to analyse a 'general' fringe pattern in order to provide solutions to the measurement problems listed above. The result of any analysis should take the form of a contiguous phase distribution. This data may be processed further to give the required measurement parameter field assuming some knowledge of the optical system and the object under test. For routine usage, the analysis should be achieved in a fully automated manner and be cost effective. Automated fringe analysis also implies that an optical system may be used by engineers of different disciplines without re-training in a new field.

The quantity of data contained in a fringe field is vast, of the order of $\frac{1}{2}$ GByte. Furthermore, the phase data required is encoded as an intensity modulation, usually a cosinusoidal or Bessel function. Image sensors, such as photodiodes or charge coupled devices, detect the intensity of the incident light. Hence some means of data decoding is required to extract the phase data.

Several stages of processing are required to overcome the problems of automatic analysis and extraction of the phase data.

i) Firstly, optical techniques are required to form the fringe fields so that the phase data can be decoded.

ii) A sensor is needed to detect the intensity in the fringe field. This requires interfacing to an analyser in order to extract the phase information. This data must then be stored.

iii) The complete data analysis system requires automation to ensure cost effectiveness.

The current solutions to these three points are described in sections 1.1 1.2 1.3.

Most of the examples in this work are of fringe fields with a cosinusoidal intensity profile. This covers a wide range of applications : static deformations, vibration analysis, compressible flows, and the measurement of temperature fields. Vibration analysis, compressible flows, and the measurement of temperature fields were stud-

ied using a pulsed laser, which also facilitates the study of transient events. The occurrence of Bessel function fringes arises from the study of vibrations when the exposure time for the interferogram is much longer than the temporal period of the motion [18].

1.1 Interference Phase Determination

Optical techniques have been developed since the first use of optical interferometers [19]. Subsequently moire systems, holographic interferometers, electronic speckle pattern interferometers (ESPI), and shearing systems have been used. In the case of a static deformation applied to an object, the fringes formed are cosinusoidal. This can be represented by a spatially varying intensity, $i(x,y)$, given by :

$$i(x,y) = a(x,y).(1 + m(x,y).\cos(\phi(x,y))), \tag{1.1}$$

where $a(x,y)$ is the local background intensity, $m(x,y)$ is the modulation of the fringes, and ϕ is the phase change to be measured.

The initial analysis of these fringes was by 'fringe counting' (also called the allocation of fringe order numbers); either directly from the fringe field or a photographic record. The amount of information contained in the field was reduced by considering only the centres of the bright (or dark) bands making up the fringe pattern. As a result phase contours were produced, and not a contiguous phase map.

A single cosinusoidal fringe map contains insufficient data for a unique solution of the phase to be determined. This is due to the ambiguous nature of the fringes. The point may be illustrated with reference to figures 1.1 1.2 1.3. Figure 1.1 shows an original phase function to be recorded. An interferometer produces the cos of this function (see figure 1.2). This graph may be interpreted in many different ways as the phase gradient is not encoded into the cosinusoidal fringe pattern. Hence at the fringe maxima, it is unknown whether to add or subtract 2π from the phase. The possible interpretations of figure 1.2 are shown in figure 1.3.

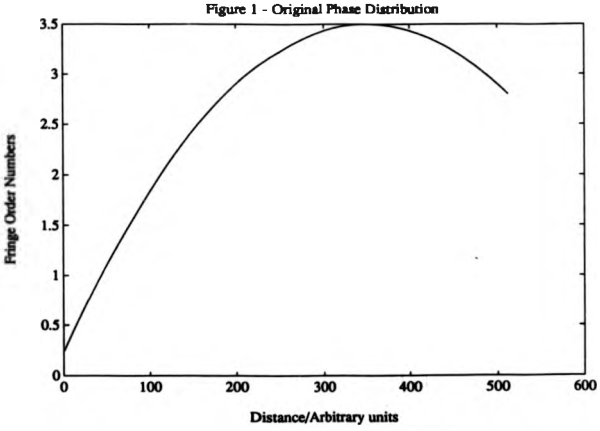


Figure 1.1: Original Phase Distribution to be Recorded

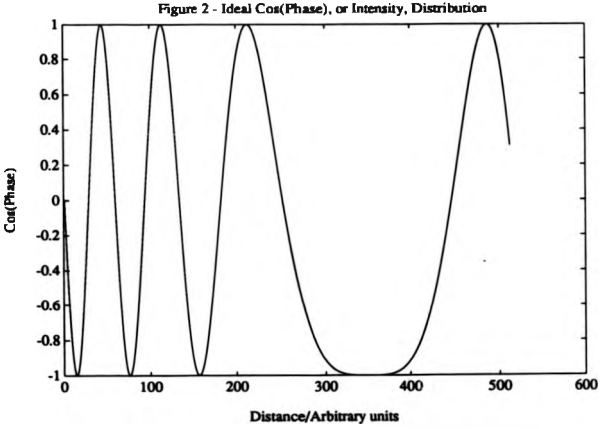


Figure 1.2: Intensity Distribution Corresponding to the Phase Distribution

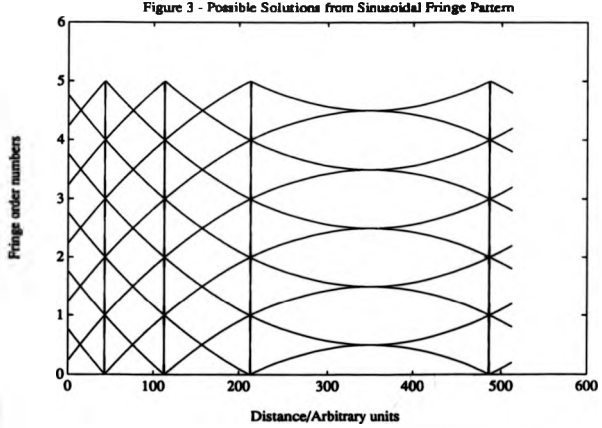


Figure 1.3: Possible Interpretations of Cosinusoidal Fringe Pattern

Fringe counting relies on human interpretation of the fringe pattern. In this way the sign of each fringe may be designated from one fringe centre to the next. This process requires *a priori* information about the experiment and the expected result. Even with this knowledge of the field, there is no guarantee that the correct phase distribution will be achieved. For the successful use of fringe counting, the user requires detailed knowledge of both the optics and the item under test.

A higher measurement resolution could be achieved by interpolation between fringe centres. This technique allows measurements to $\frac{1}{50}$ of a fringe to be made in a few cases [20, section 2.3] [19].

1.1.1 Optical Heterodyne Analysis

The limitations described above were largely removed using the heterodyne and quasi-heterodyne techniques [21, 22, 23]. By optical heterodyning the phase information could be directly extracted from the field. This was achieved by reconstructing the fringe data with a beat frequency superposed on the optical frequency. In practice this is achieved by forming the fringe field with two waves of slightly different optical frequency [21]. These waves w_1 and w_2 may be represented by :

$$w_1(x,y) = a_1(x,y)\Re e^{i[\omega_1 t + \phi_1(x,y)]}, \tag{1.2}$$

and

$$w_2(x,y) = a_2(x,y)\Re e^{i[\omega_2 t + \phi_2(x,y)]}, \tag{1.3}$$

where $a_i, i = 1,2$ is the amplitude of each wave, $\omega_i, i = 1,2$ is the optical frequency, $\phi_i, i = 1,2$ is the phase of the wave, and $\Re$ denotes the real part. A photodetector will measure the time-dependent intensity of the superposition of these two waves [21] :

$$i(x,y,t) = |w_1 + w_2|^2 \tag{1.4}$$

$$i(x,y,t) = a_1{}^2(x,y) + a_2{}^2(x,y) + 2a_1(x,y)a_2(x,y)\cos\left[(\omega_1 - \omega_2)t + \phi_1(x,y) - \phi_2(x,y)\right] \tag{1.5}$$

$$i(x,y,t) = a_1{}^2(x,y) + a_2{}^2(x,y) + 2a_1(x,y)a_2(x,y)\cos\left[\Omega t + \phi_i(x,y)\right] \tag{1.6}$$

where $\Omega = \omega_1 - \omega_2$ and $\phi(x,y) = \phi_1(x,y) - \phi_2(x,y)$. Providing that Ω may be resolved by the photodetector, the optical interference phase can be measured electronically as the phase of the beat frequency. Once the phase is available rather than a function of the phase, the ambiguity in assigning fringe orders is removed. Thus heterodyning allows a further stage of the analysis process to be automated.

Two disadvantages remain for heterodyne analysis. Firstly, the data may only be read out in a point-wise manner. This inherently limits the processing speed for a

complete interferogram. Secondly, due to the trigonometric nature of the fringes, a given phase difference δ cannot be distinguished between $\delta + m2\pi$, m integer. To overcome this, a count of the multiples of 2π is also made as the data is scanned out. This count can be made unambiguously, as opposed to fringe counting, as the interference phase itself is known at all points.

1.1.2 Quasi-Heterodyne Analysis

The quasi-heterodyne technique allows the interference phase to be calculated unambiguously as in heterodyne analysis. In this case, the intensity of a complete fringe field is measured simultaneously with a detector array. Several fringe images are acquired where the phase of the fringe field is varied either in discrete steps [22] or continuously [24]. When discrete steps are added to the fringe field phase, the method is usually referred to as 'phase stepping'. A number of intensity values are then available for each point in the fringe field. From this data the interference phase can be calculated in a point-wise manner.

Conceptually, the minimum number of images required for quasi-heterodyne analysis is two [25]. In this case, if a positive phase shift of $\frac{\pi}{2}$ is applied the direction in which the fringe maxima move indicates whether the fringe order number is increasing or decreasing. In practice three, four, or five images are normally used [23, 8, 26]. When three images are used with a phase step of 90 degrees, the three images take the form [23] :

$$i_0(x,y) = a(x,y).(1 + m(x,y).\cos(\phi(x,y))), \tag{1.7}$$

$$i_{90}(x,y) = a(x,y).(1 - m(x,y).\sin(\phi(x,y))), \tag{1.8}$$

$$i_{180}(x,y) = a(x,y).(1 - m(x,y).\cos(\phi(x,y))), \tag{1.9}$$

where $i_\theta(x,y)$ is the intensity at the position (x,y) when the phase step is θ. This can be considered as a set of simultaneous equations with unknowns ϕ, $a(x,y)$, and $m(x,y)$. By algebraically eliminating $a(x,y)$ and $m(x,y)$, the phase may be

calculated by :

$$\phi(x,y) = \tan^{-1}\left\{\frac{2i_{90}(x,y) - i_0(x,y) - i_{180}(x,y)}{i_0(x,y) - i_{180}(x,y)}\right\}. \qquad (1.10)$$

The same expression can also be derived as a special case of the analysis by Grievenkamp [27], where the least squares fit of a sinusoid to the phase stepped intensity values is considered.

The phase values are evaluated using an arc tangent function [23]. Hence the results are 'wrapped' into the range $-\pi$ to π. The map produced by calculating the phase across the whole field is called a wrapped phase map. For a total variation in optical path difference across an image of $> 2\pi$, the wrapped phase map will contain discontinuities. This makes the whole field difficult to interpret and can cause errors when further computation must be performed on the data, for example the determination of strain. To overcome this problem, the phase discontinuities must be offset by 'phase unwrapping' [22]. In this process 2π is added or subtracted from the wrapped phase values every time a wrapover point is reached. Whence, the unwrapped phase map is produced, containing a continuous phase distribution.

The variation of background illumination and modulation depth may also be calculated using :

$$a(x,y)m(x,y) = \frac{i_0(x,y) - i_{90}(x,y)}{\cos(\phi(x,y)) - \sin(\phi(x,y))} \qquad (1.11)$$

and

$$a(x,y) = i_0(x,y) - a(x,y)m(x,y)\cos(\phi(x,y)). \qquad (1.12)$$

Variations in these two parameters allow the data quality to be assessed across an image. The fringe contrast, or modulation depth, $m(x,y)$ also allows object discontinuities such as holes to be detected automatically [28, 29].

The use of four images in the phase stepping algorithm allows the phase step to be determined without calibrating the phase shifting mechanism. Alternatively, an increase in the number of images used can introduce a degree of redundancy in the

solution which in turn minimises the errors in the calculated phase (if there exists errors in the input parameters). For example, Hariharan proposed using five images with a phase step of 90 degrees to minimise measurement errors due to inaccurate phase steps [26].

With the current commercial availability of detector arrays and image digitisers, this method offers rapid data acquisition and direct computer analysis.

1.1.3 Carrier Fringe Methods

Fringe order ambiguities can also be removed by the addition of 'carrier fringes'. In this technique, a linear phase variation is added to the fringe field as a function of image co-ordinate. This can be achieved in practice by tilting one of the mirrors in an interferometer, thereby giving a 'tilted wavefront'. The intensity in the fringe pattern may then be represented by :

$$i(x,y) = a(x,y).(1 + m(x,y).\cos(2\pi f_o x + \phi(x,y))), \qquad (1.13)$$

where f_o is the spatial carrier frequency assumed to be in the x direction. The carrier frequency must be sufficiently high so as to produce monotonically increasing fringe order numbers. When this is achieved, the field can also be described as a finite fringe field, and is characterised by the absence of any closed loop fringes. An example of such a field is given in figure 1.5 (contained in section 1.3.1), which was produced by a classical specular interferometer.

A finite fringe field can be processed by fringe counting without ambiguity as the fringe orders are known to increase monotonically in the direction of the tilted wavefront. The phase data obtained represents the required measurement parameter added to the phase of the tilted wavefront. The phase change due to the tilt must therefore be calculated and subtracted from the original phase information obtained. The required parameter may then be calculated from the resulting phase distribution.

The usefulness of carrier fringe methods in interferometry was not fully realised until electronic imaging hardware and more powerful computer systems became avail-

able 1.3.5.

1.2 Hardware Requirements for Fringe Analysis

The most obvious means to handle large quantities of information whilst allowing flexible means of data processing is to utilise computer technology. Computer technology has developed rapidly over the past 20 years and with ever-increasing processing power it has enabled the practical implementation of many of the optical techniques described above. The general requirements of a computer system to perform fringe analysis are listed below.

i) Computational power. This is limited by the processing speed of the microprocessor.

ii) The facility to run high level computer languages hence allowing flexible user programing.

iii) Large data storage space.

iv) The facility to interface image digitisers, and other devices, to computers thereby forming an integrated system for industrial use.

v) Rapid data transfer from the storage devices and computer interfaces.

vi) A suitable means to output the data.

vii) Portability, for on-site industrial inspection and analysis.

The requirements for each of these items is continually growing due to the industrial need to evaluate more complex problems.

To take advantage of flexible computer analysis, the data must first be interfaced to the computer system. This involves two components : a sensor to detect the intensity in the fringe field, and a digitiser to convert the output of the sensor into a digital form.

Initially, computers were used in conjunction with a digitising tablet to allow the entry of fringe maxima or minima from a photograph. In this case, the sensor was the human eye and discrete x, y co-ordinate pairs were input to the computer. The data was detected and processed in a sequential manner, thereby limiting the processing speed. Some increase in the processing rate was achieved by replacing the human eye with a photodiode and computer interface [15]. The computer utilised the detected intensity to scan the diode along fringe maxima or minima. These methods of acquiring digital fringe information were time consuming and only provided phase data at discrete points along the fringe centres.

Electronic cameras and image digitisers were needed to form a more efficient interface between the optical and computer systems. In this way, a grid of intensity values is detected simultaneously and some of the problems described above are removed. A detector array inherently possesses a parallel detection of the intensity at a set of regularly spaced grid points. The limiting factors are then the size of the array and the speed at which the data can be read out from the individual sensors. One of the first reports of such a device being used in fringe analysis was made in 1974 by Bruning et al [22]. The output from a two dimensional array of 32 by 32 photodiodes was digitised electronically and fed directly into a computer. The computer controlled the image digitisation process and the operation of a piezo-electric (PZT) phase stepping device. Thus full synchronisation of phase stepped fringe field capture was possible. The quasi- heterodyne phase calculation algorithms were implemented in software to produce the phase map. This work formed the basis for many quasi-heterodyne analysis systems.

The main limitations evident from the work by Bruning et al are described below.

i) The resolution of the detector array (and digitiser) limited the range of fringe fields which could be analysed. This is determined by the Nyquist sampling theorem which states that one fringe must be imaged over at least two detector elements [30, p.67-68].

ii) An execution time of 1 minute was achieved from the Bruning system. Considerably greater computer power would be required to solve higher resolution data of general objects.

iii) The fringe fields analysed had a low noise content, hence simplifying the computational analysis. This resulted from the use of a Twyman-Green interferometer with the fringe field being projected directly onto the face of the detector array. Hence the problems of speckle noise were not present.

iv) The objects examined were of simple form, again simplifying the analysis.

Points one and two are inter-related. With the increase in resolution of electronic sensors and image digitisers the data space is augmented. Hence faster processors are required.

Progress on the automatic analysis of fringe fields was inhibited by the lack of commercial detector arrays of sufficient resolution and the shortage of sufficient computing power to perform the analysis. Advances in imaging hardware occurred with the emergence of video technology. This took the form of $\sim 600 \times 525$ pixel solid state charge coupled device (CCD) cameras. The development of real time video digitisers is technologically linked to that of computer hardware. In the mid 1980s, both the Intel 80286 range of personal computers (P.C.) and $512 \times 512 \times 8$ bit frame grabbers became available. This scale of image resolution proved sufficient to analyse many industrial problems and the computational power produced an acceptable execution speed. At this time, it became practical to attempt fully automatic processing systems for a wide variety of fringe patterns.

New advances in VLSI (very large scale integrated circuits) technology can be expected in the future as a consequence of the rapid growth that has occurred in personal computer power. Figure 1.4 shows the development of the microprocessors used in IBM and compatible P.C.'s in terms of the processor power in MIPS (Millions of Instructions Per Second) against the introduction date of the chip. This graph shows an increasing rate of growth especially with the introduction of the i860 pro-

cessor. The increased processing power facilitates either quicker execution times or the implementation of more complex software analysis schemes. An example of this is the application of artificial neural networks (ANN's) which have the capacity to learn from previous experience. This type of 'intelligent' software may be applicable to the analysis of general fringe patterns in the future. Similar advances in CCD resolution are currently being achieved with commercial systems for 2048 x 2048 pixels reaching the marketplace (although at an inhibitive price for fringe analysis systems). The increased image resolution means that fringe fields with higher spatial frequency fringes will be adequately sampled and processed. The combination of these two technologies will provide greater flexibility in the range of applications for the computerised, quantitative analysis of interferometric fringe data.

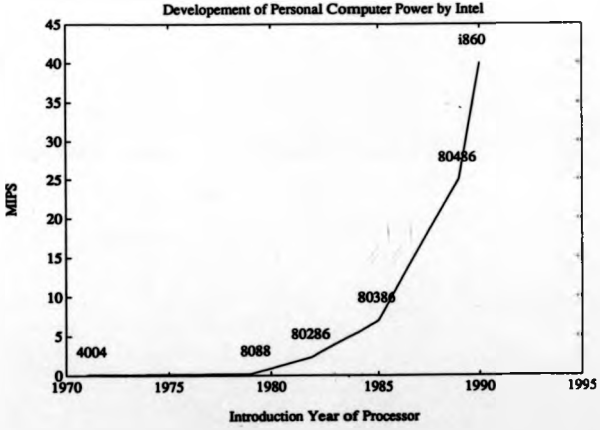


Figure 1.4: Developments in Personal Computer Processing Power

1.3 Automated Evaluation of General Fringe Patterns

With the availability of computer hardware and high level programming languages, the theoretical ideas and algorithms for fringe analysis could be implemented. This allows flexible system development and the incorporation of new ideas in software. The complexity of the analysis is then limited by the algorithm used. Furthermore, the automation of a fringe analysis system from the operation of peripheral devices to image capture, data processing and storage may all be controlled by means of software.

The objective of software analysis of fringe patterns is then to perform fully automatic processing of the fringe data resulting in a contiguous phase distribution. At this stage, it is worth summarising the problems facing an automatic fringe analysis package. This list has been compiled with reference to work by Trolinger [31] and other workers [32, 33].

A–i) speckle noise,

A–ii) diffraction noise,

A–iii) uneven fringe background,

A–iv) varying fringe contrast,

B–i) unknown fringe sign,

C–i) discontinuous and split fringes,

C–ii) closely spaced fringes,

C–iii) broad cloudlike fringes,

C–iv) extraneous fringes,

D–i) regions blocked by opaque objects,

D–ii) lack of a known reference position,

E–i) the data reduction process should be fast and provide sufficient resolution of phase and the spatial detection of the phase.

These points may be grouped into several categories. Items A–1 to A–4 are caused by the scattering characteristics of the surface under examination and any non-uniformity in the interfering wavefronts. These problems are normally approached by pre-processing the fringe fields to attain better and more uniform signal to noise ratio (SNR) over the entire field [32, 34].

The unknown fringe sign (item B–1) may be reconciled by the optical method used (heterodyne, quasi-heterodyne, carrier fringe etc.) [35], or by prior knowledge of the experiment and human interaction (fringe counting). The latter is a non-automatic process.

Items C–1 to C–4 specify the degree of complexity of the phase data in the field. Problems may arise in the correct determination of the interference phase for several reasons : when the phase changes rapidly and the fringe density approaches the Nyquist sampling limit of the detector [33], or the fringes are very broad and some noise is present in the data. Other ambiguities may also be present such as discontinuous fringes. In this case different answers will be obtained depending on the route taken to relate the phase at one point to that at another [36, 37, 38, 39]. Extraneous fringes are due to extra coherent waves being detected which add faint fringes to the desired interference pattern.

The two points, D–1 and D–2, are partly caused by the experimental arrangement and the 'view' of the object which is obtained. A point in the field which is known to have zero phase change, or equivalently no movement, allows absolute measurements to be made. This is not always a necessary requirement, for example in NDT it is often the relative phase change across a specimen that is of interest.

Finally, fast data processing and a specified resolution is required for the measurement method to be industrially satisfactory.

For fully automated analysis, all the above points must be considered. Some

of the problems are solved by the use of the appropriate optical methods, e.g. the determination of fringe order numbers. The extent to which the data quality is improved and the phase determined is dependent on the complexity of the algorithm and the initial quality of the data. A further consideration is the analysis speed which is especially relevant to the industrial application of the system.

1.3.1 Pre-Processing of Fringe Data

The extent to which pre-processing must be applied is dependent on the optical method used to generate the data. For fringe fields generated by classical specular interferometry and some holographic fringe patterns the data may be of sufficient quality for direct phase calculation. However this cannot be guaranteed, and the case of speckle interferometry and moire systems need to be considered. Examples of interferometric, holographic, and ESPI fringe fields are given in figures 1.5 1.6 1.7. In each case the fringes are cosinusoidal and have been captured using a CCD camera and $512 \times 512 \times 8$ bit frame grabber. It is clear from the form of these images that the noise content in each case is very different. The noise may be assessed further by calculating the fast Fourier transform (FFT) of a row of pixels in the image. The resulting amplitude spectra are shown in figure 1.8. The signal power in the high frequency range can be used as an estimate of the noise content. The proportion of noise is seen to increase from interferometric to holographic to ESPI fringe images.

The methods of pre-processing fringe data have been reviewed as part of a paper by Reid on automatic fringe analysis[32]. Further details on the methods are contained in references [33, 6]. It is assumed that the fringe data has been digitised by a suitable camera and frame grabber.

The pre-processing technique to be applied depends on the type of noise present. When the spatial frequency of the noise is higher than that of the fringe data then a low pass filter may be applied. This normally takes the form of a window (or kernel), with either the average or median of the intensities contained within the window being determined. This value is then used to supplant the intensity at the

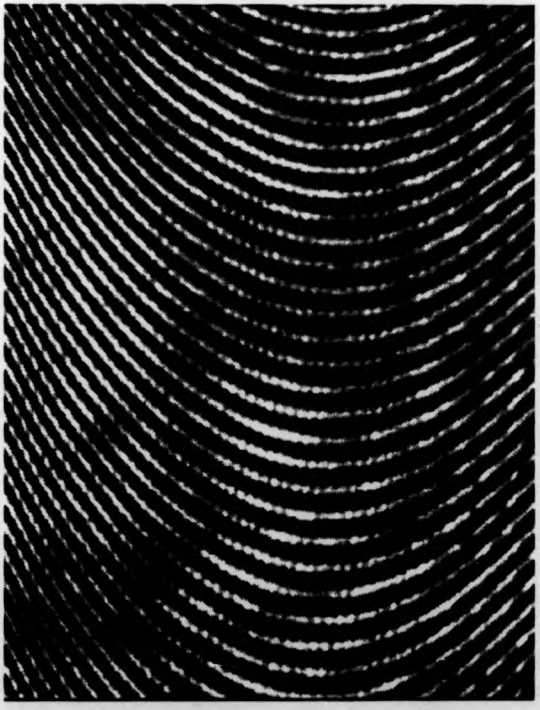


Figure 1.5: Example of an Interferometric Fringe Pattern

Figure 1.6: Example of a Holographic Fringe Pattern

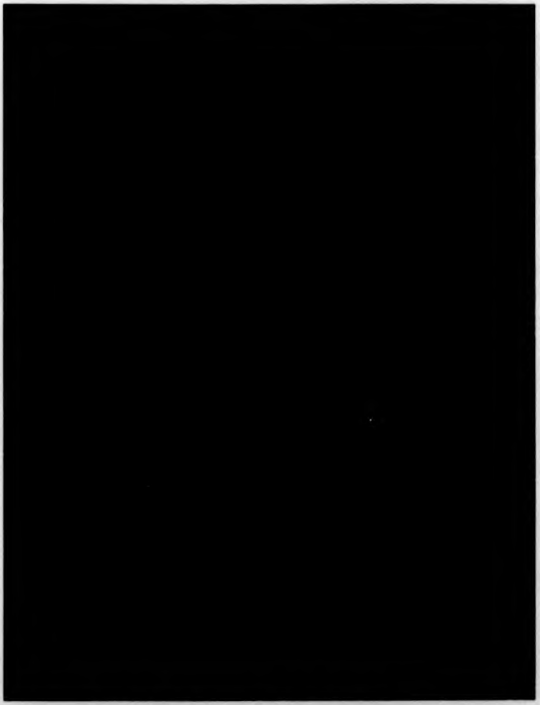


Figure 1.7: Example of an ESPI Fringe Pattern

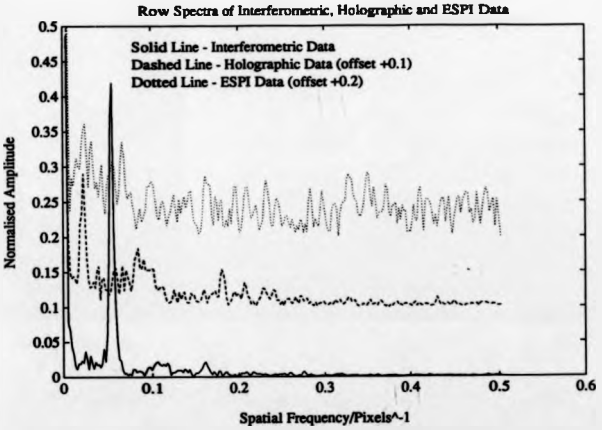


Figure 1.8: Amplitude Spectra of Interferometric, Holographic, and ESPI Fringe Data

centre of the window. The window shape is normally a square array of pixels or a cross. More sophisticated approaches have also been implemented, for example the spin filter algorithms of Yu [40]. This technique uses the average grey scale over a line in a specific direction. The line is rotated, spun, around the point of interest until it lies as close as possible to a constant phase line. The average value over the line is then assigned to the point in question. In this way, a more refined filtering of the data is produced which reduces the blurring effect of conventional square window filters. Square shaped windows remain the most commonly used because of the rapid execution times, approximately 30 times faster than the spin filter algorithm.

An effective reduction in speckle and electronic noise is produced by all the methods described. This can be illustrated on the ESPI image in figure 1.7. By applying four successive iterations of an averaging filter over a 3×3 window, the image in figure 1.9 is produced. The spectra of the same row in the original and the filtered images are shown in figure 1.10. The noise reduction achieved is apparent.

Similarly, in the case of uneven illumination or fringe contrast, the fringe data may be separated from the background by a high pass filter. This approach was implemented by Becker [41] using regional averages. These values are smoothed over the whole image and the result subtracted from the interferometric data to form a more uniform background intensity.

The algorithms above may also be implemented in frequency space rather than in the spatial domain [42, 6]. This is achieved by evaluating the FFT of a raster in the fringe data. A mask may be placed in the fourier plane and then the inverse transform taken to yield modified intensity data. By appropriate choice of the mask cut-off frequencies, the desired noise reduction, high frequency or background variation, can be achieved on the fringes [42].

One disadvantage of the frequency domain approach is the execution time which can be ~ 20 times greater when compared with the corresponding spatial processing method. This can be explained by comparing the actual filtering processes being performed in each case. The FFT method produces the convolution of the window function with the image, whereas the spatial domain approach only considers a small

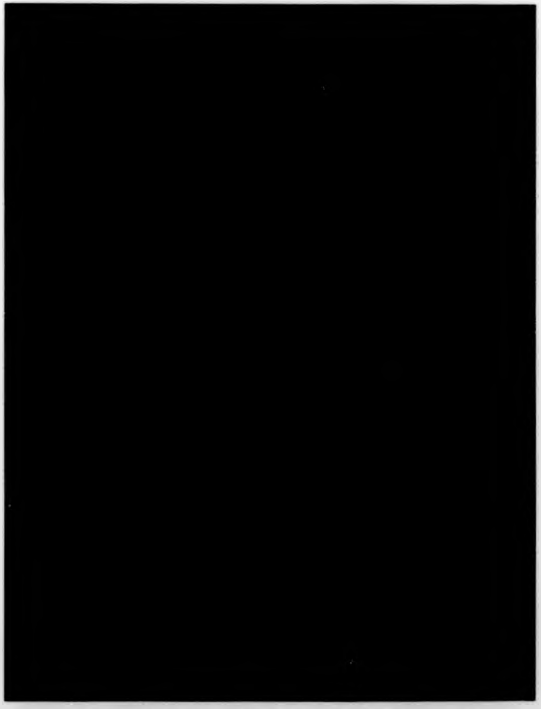


Figure 1.9: ESPI Fringe Image After 4 Passes of a 3*3 Averaging Filter

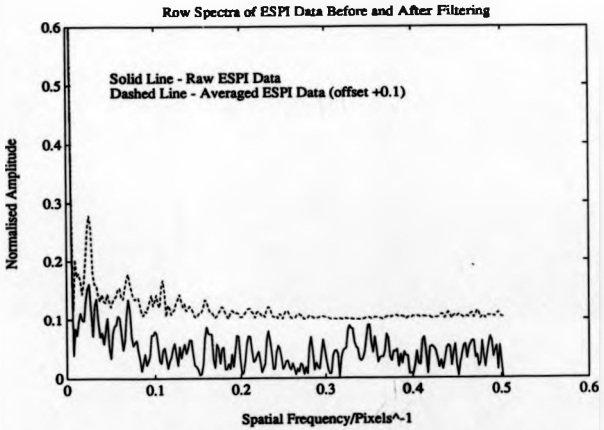


Figure 1.10: Comparison of ESPI Amplitude Spectra Before and After Filtering

region of the field (3 x 3 pixels) with unity amplitude for each point in the window. The time penalty in using frequency domain filtering has recently been overcome by using dedicated hardware [43].

The noise reduction algorithms as described above have a limited usefulness when the noise source has approximately the same frequency as the fringes. In this case it is very difficult to improve the data effectively. When the noise is stationary, i.e. not related to the position of the fringes, it is possible to subtract the background variation by using two fringe images with a phase difference of π radians between them [31, 44]. The two images then represent the same fringe field but a bright fringe in one image becomes a dark fringe in the other. By point-wise subtraction of the two fields, the noise will be cancelled out. The process produces higher contrast fringes at the expense of a more complex optical arrangement.

In all the references cited in this section, the noise reduction algorithm has been suited to the data concerned. Hence the analysis is automatic only if fringe patterns of a similar quality are to be processed. A topic of current research is the automatic detection of the necessary degree of smoothing.

1.3.2 Fringe Tracking

Some of the first codes to really tackle automatic fringe processing were based on the idea of 'fringe tracking'. The aim of this technique is to use an algorithm to automatically track the fringe maxima (or minima), see for example [45, 46, 47, 48, 49, 32]. The fundamental concept is to examine a window of pixels around a given point. The average intensity is computed for all possible directions leaving the point in question, see figure 1.11. To prevent the track from continually circling, only three of these directions from those possible are normally considered : the original direction to reach the point in question, and the directions either side of the original direction [45]. If a fringe minima is to be traced, the minimum of the average intensities along these three directions is taken to reach a new point. The process is repeated until some criteria is satisfied or the edge of the image is met.

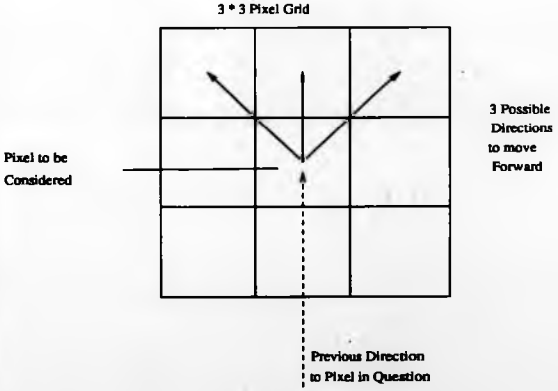


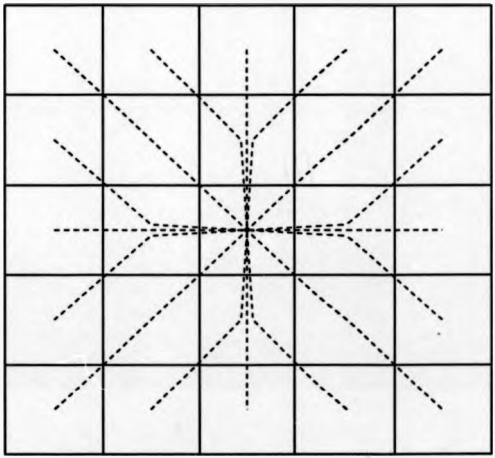
Figure 1.11: Possible Directions to Track Fringes

Other methods based on skeletonising the fringe data before fringe tracking have been used [50, 51, 15]. Skeletonising normally results in a binary image achieved by the application of a threshold operator. The 'white' fringes in the image can then be thinned and tracked to form continuous lines representing fringe maxima. The analysis of a combustion flame interferogram by Bryanston-Cross et al [15] is interesting as a spatial resolution of at least 4000 x 4000 pixels was required to adequately sample the fringe field. This work demonstrates that fringe tracking may be performed over a much larger data space than the standard video resolution of 512 x 512 pixels.

An implementation of the algorithm by Button [45] has been made with a 5 x 5 pixel window. The algorithm has been extended by considering the 16 directions shown in figure 1.12 and the automatic detection of a starting point and initial direction for tracking. The result for five fringes of the interferogram in figure 1.5 is shown in figure 1.13. A similar result can be obtained for the holographic fringe data in figure 1.6, four tracked fringes are shown in figure 1.14. In this case the fringe data required filtering by a single pass of a 3 x 3 averaging window prior to fringe tracking. The starting points of the tracked fringes were set at just below the start of the fringe data in the image. The need for filtering can be demonstrated by examining figure 1.15 which shows the result when the same tracking algorithm is executed on the raw fringe data. This implies that a certain image quality is required for successful fringe tracking.

Due to these problems, the development of fringe tracking algorithms have followed two separate paths : either the invention of more sophisticated algorithms for both image enhancement and fringe tracking, or implementing semi- automatic systems requiring operator intervention. Both these approaches have been followed by different groups.

Yatagai et al [46] implemented a more critical test for determining fringe maxima points by testing in several directions for each point. Based on the statistics of speckle noise [20, p.30-36], it is more likely that a dark speckle will occur in a bright fringe than a bright speckle in a dark fringe. Hence it is more difficult to track fringe maxima than minima in the presence of speckle noise and the more sophisticated tests are



5 * 5 Pixel box for Determining Tracking Direction

---------- Directions over which average of Pixel Grey Scales is Taken

Figure 1.12: 5 * 5 Window for Implementing Fringe Tracking

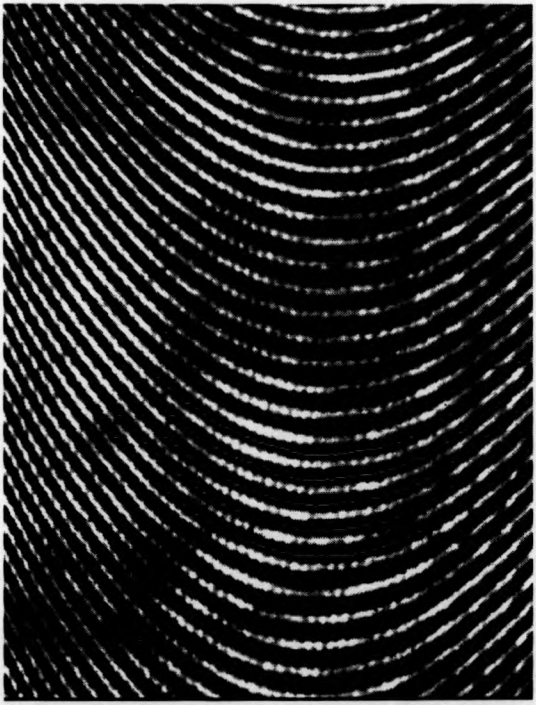


Figure 1.13: Five Tracked Fringes in an Interference Fringe Pattern

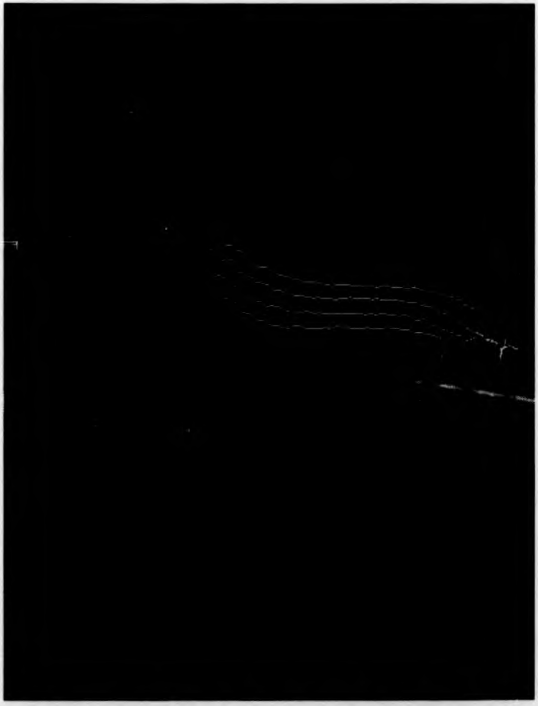


Figure 1.14: Four Correctly Tracked Fringes in a Filtered Holographic Fringe Image

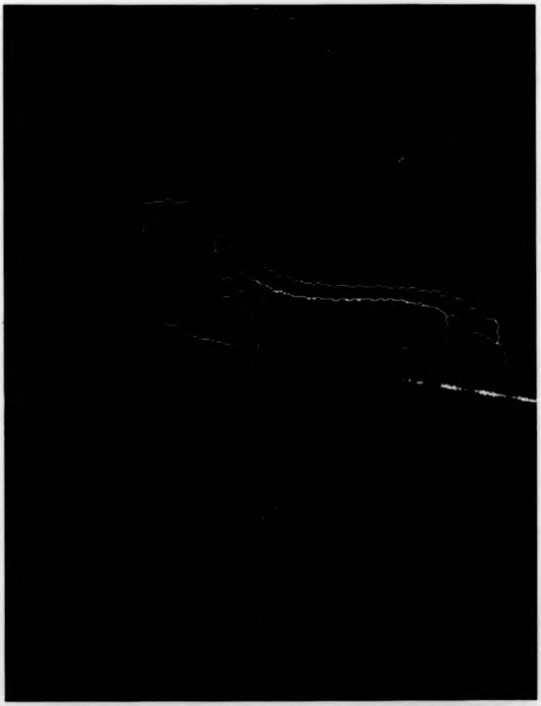


Figure 1.15: Four Fringes Incorrectly Tracked in a Raw Holographic Fringe Image

necessary.

The problem of broken fringe tracks has been addressed by Liu et al [51]. In this work the ends of fringe tracks which are not at the edge of the image are identified. A sector is then defined using one such fringe end point as the centre. If a second fringe end point is found within the sector then the two end points are joined.

In contrast Funnell [47], Yatagai et al [46] and Parthiban [52] have developed interactive fringe tracking systems. Their philosophy concedes that a fully automatic algorithm is not possible and consequently a human interface is required to correct for mistakes.

Assuming that the fringe extrema can be tracked correctly, there are still two disadvantages which are apparent for all fringe tracking methods. Firstly, the fringe order numbers (point B–1) must be assigned manually unless some prior knowledge can be utilised, or carrier fringes have been added to the original fringe field. Secondly, phase data is only available on the fringe maxima and minima thereby restricting the use of the data especially for strain analysis where spatial derivatives are required. This may be overcome by fitting a curve to the known fringe maxima and/or minima points, see for example [53, 54]. However this analysis is not usually employed due to the execution time required [32].

Fringe tracking cannot therefore form a full solution to the automatic analysis of interferograms, see section 1.3. However, the method is still used in classical lens testing interferometers because of the simplicity of the optical system required, and the high performance of fringe tracking algorithms on specular interferometric data (see figure 1.13). A fringe tracking analysis system has some advantages in certain other cases.

i) Where an infrequent analysis is to be performed.

ii) The complexity of heterodyning optical systems is not practical.

iii) In the analysis of time average fringe fields (containing Bessel function fringes).

iv) When a single record exists of a non-repeatable event.

1.3.3 Automation of Heterodyne Phase Analysis

The first automatic holographic heterodyne system was constructed by Dändliker and can produce data at a rate of approximately 1 second per point in the interferogram [28]. By scanning the detector using a motorised x, y stage, a contiguous phase distribution can be obtained for the whole field at any required spatial resolution. The data collection and detector position are computer controlled. The limit of 1 second per point arises due to the re- positioning of the scanned detector and the time to log the detected phase data. It may be possible to increase the data collection rate by continuously scanning a detector across the fringe field and reading the phase data at intervals providing that the stability of the optical system can be maintained. A different approach was adopted by Massie [55] where an image dissector camera was used to select which part of the interference field is sampled at any point in time. The system constructed allowed data collection at $5\mu s$ per point. Therefore, to cover the number of points available in a 512×512 framestore, i.e. 262144, requires approximately 1.3 seconds.

The advantages of heterodyne phase measurement are the very small phase differences which may be resolved unambiguously. Up to $\frac{1}{1000}$ of a fringe can be achieved. The spatial resolution is high and only limited by the motorised traverse stage (or image dissector camera) and the characteristic speckle size in the image. Some noise immunity is also gained as the phase to be measured is independent of extraneous fringes (item C–4) which are not modulated at the beat frequency, and the interfering wavefront amplitudes (items A–3 and A– 4, see equation 1.6).

Despite these benefits, heterodyne systems are confined to a stabilised laboratory bench [56]. This is due to the stability requirements of the reconstruction optics over the duration of the data read-out and the need for accurate hologram repositioning after development. Clearly this is not appropriate for performing rapid measurements in an industrial environment. There is no facility to pre-process the fringe data (points A–1 to A–4) and hence successful analysis is dependent on achieving sufficient signal in the detected intensity at the recording stage of the interferogram [57].

1.3.4 Automatic Quasi-Heterodyne Fringe Analysis

The automation of quasi-heterodyne analysis systems was apparent after the work by Bruning et al. In general a computer system is linked to : a CCD video camera via an image frame grabber, and a device for varying the phase of the fringe field. Synchronisation of the image capture system and fringe field phase is achieved through software. It has been shown by subsequent workers that the set of phase stepped images (typically three, four, or five, see section 1.1.2) can be captured in consecutive video frames (see for example [8]). This is done by moving the phase shifting device in the read-out time of the CCD camera. Hence digital image capture requires $\frac{n}{25}$ of a second, where n is the number of images.

Once the data is in digital form, the data processing may be achieved by a sequence of software functions. This allows flexibility in the analysis and independent fringe pre-processing (section 1.3.1) followed by phase computation. Many phase calculation algorithms are available and have been reviewed by Creath and Wykes [58, 56]. As described in section 1.1.2, the resultant phase is evaluated modulo 2π. The image formed is called a 'wrapped' phase map. The wrapped phase map corresponding to figure 1.6 is shown in figure 1.16 where the phase values between $-\pi$ and π have been encoded as grey scales from 0, black, to 255, white. The cosinusoidal images were smoothed using a 3×3 averaging filter prior to calculation of the wrapped phase values.

The automatic processing of a wrapped phase map, i.e. 'phase unwrapping', will be discussed after the following section.

The advantages of quasi-heterodyne analysis are :

i) the rapid, automatic data acquisition which directly yields the fringe data in a digital form,

ii) and the facility for software analysis.

The phase measurement resolution is also high, up to $\frac{1}{100}$ of a fringe [58], whilst the spatial resolution is fixed by : the number of pixels in the digital frame grabber (or

Figure 1.16: Wrapped Phase map Evaluated by Quasi-Heterodyne Analysis

CCD camera whichever is the smaller), the image magnification, and the Nyquist sampling limit. In practice, 512 x 512 pixel resolution is sufficient for many problems on moderately sized solid surfaces. The application to phase objects is restricted to simple examples owing to the high spatial frequency of the fringes which is often found. Some extensions are possible by breaking the image down into regions and solving each one separately. The data from the individual sections may then be merged together.

1.3.5 Automated Carrier Fringe Analysis

An interferogram with sufficient carrier fringes gives a monotonic increase in fringe order numbers, see section 1.1.3. This may be utilised by inspecting the frequency spectra of a raster which is parallel to the direction of the tilted wavefront, the x direction in equation 1.13. The fringe pattern formed by a standard interferometer, figure 1.5, can be considered as a carrier fringe image with monotonically increasing fringe orders. The complete spectra of an image raster is seen in figure 1.17. This is characterised by a zero order peak and symmetrical side lobes for positive and negative spatial frequencies (the vertical scale of the figure has been limited to aid interpretation of the spectra).

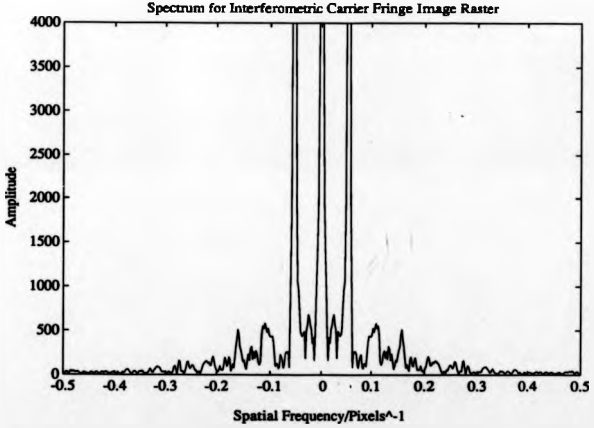


Figure 1.17: Symmetrical Spectra of Carrier Fringe Image Raster

The phase variation to be measured exists as a broadening of the side lobes in

frequency space. The terms may be recognised mathematically by re-writing equation 1.13 as :

$$i(x,y) = a(x,y) + b(x,y).\cos(2\pi f_o x + \phi(x,y)), \tag{1.14}$$

$$= a(x,y) + c(x,y).e^{i2\pi f_o x} + c^*(x,y).e^{-i2\pi f_o x}, \tag{1.15}$$

where

$$b(x,y) = a(x,y).m(x,y), \tag{1.16}$$

and

$$c(x,y) = \frac{1}{2}b(x,y).e^{i\phi(x,y)}, \tag{1.17}$$

f_o is the carrier frequency, and * indicates the complex conjugate. If the Fourier transform is applied to equation 1.15 with respect to x then

$$I(f,y) = A(f,y) + C(f - f_o, y) + C^*(f + f_o, y), \tag{1.18}$$

where the capital letters denote the Fourier spectra and f is a spatial frequency variable in the x direction. The three terms of equation 1.18 are separated in the frequency domain providing that the spatial variations of $a(x,y)$, $b(x,y)$ and $\phi(x,y)$ are slow compared with f_o. When this condition holds the zero order peak and the two side lobes as described above are produced. This analysis method was first invented by Takeda et al [59].

The interference phase ϕ may be found by filtering the data in frequency space. One side lobe is isolated and translated to the origin producing the spectra in figure 1.18 (again the vertical scale is limited to aid interpretation). The inverse FFT is performed to obtain $c(x,y)$ which now contains both real and imaginary parts and the phase can be calculated pointwise by [60]

$$\phi(x,y) = \tan^{-1}\left\{\frac{\Im[c(x,y)]}{\Re[c(x,y)]}\right\}, \tag{1.19}$$

where $\Re$ denotes the real part and $\Im$ the imaginary part. Again, due to the arc tangent

function, the phase is determined modulo 2π and a wrapped phase map results.

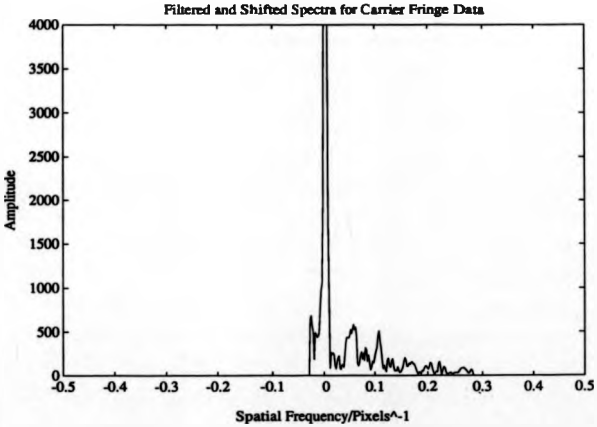


Figure 1.18: Filtered and Translated Carrier Fringe Spectra

The method is useful in that only a single image is required for the analysis as opposed to phase stepping where between two and five images are normally used. The FFT processing allows noise suppression in the frequency domain as a by- product of the phase calculation (see section 1.3.1). A suitable window function can be applied in the frequency plane to eliminate different types of noise [42, 32]. This has been specifically demonstrated by Frankowski who defines the frequency bandwidth in terms of the Nyquist frequency and the carrier frequency [61].

The spatial resolution of the measurements is again limited by the method of digitisation which is usually a video camera and digital frame grabber. The phase measurement accuracy is comparable to phase stepping at $\frac{1}{100}$ of a fringe. However

there are problems in achieving this level of precision as described by Kujawinska [62]. A window must be applied to the original data in the spatial domain to remove the effect of data discontinuities at the ends of the raster. The window also functions to make the data appear periodic and prevent leakage of energy from one frequency into the surrounding spectra. Often a Hanning or Hamming window is used. Other workers prefer the use of a $\cos^4$ window [63] to minimise leakage. A further error is due to edge effects of objects in the field which can introduce high frequency components to the frequency spectra. This has been demonstrated by Takeda and Kujawinska [59, 62] and inherently limits the measurement resolution in the neighbourhood of the edge. The problem may be minimised by extending the fringe pattern beyond the edge of the object [62]. The final problem is the accurate subtraction of the carrier frequency. If this is not done precisely, a residual ramp is left on the phase data. Several approaches to this problem are possible.

i) After filtering the data in frequency space to obtain a single side lobe, the inverse FFT can be performed without translating the side lobe to the origin [61]. A wrapped phase map is produced but still containing the carrier frequency. An image with only the tilt fringes can then be processed in the same way and subtracted from the first phase distribution.

ii) Obtain an image with only tilt fringes. Perform an FFT as before and analyse the spectrum to accurately locate the carrier frequency without the effect of the phase distribution to be measured.

iii) Utilise some knowledge of the object under test to identify rasters with zero deflection. These rasters may then serve as a reference and any phase change present be offset to zero [64]. This is clearly not an automatic process and unsatisfactory as it relies on knowledge of the object deformation which is the measurement required in the first place.

If one of the above methods is used a measurement accuracy of $\frac{1}{100}$ of a wavelength is possible within 90% of the edge of an object [62].

Other extensions to the basic method have been made. Macy [65] extended the analysis to two dimensions necessitating a two-dimensional equivalent of the window function applied. Kreis [60, 42] shows that the method can be modified to analyse fringe data when no carrier fringes are added. In the analysis, two wrapped phase maps are formed by using different filters in the frequency domain. Firstly, the vertical frequencies are restricted to varying in a positive manner, and secondly, the horizontal frequencies. The operator then has to combine the two wrapped maps in the regions where each one correctly indicates the phase variation. This cannot be done unambiguously without prior knowledge of the expected phase distribution. Therefore this is not an automatic process. Kreis suggests the use of two images without carrier fringes but with a phase step between them in order to resolve the ambiguity and thereby construct the wrapped phase map [60].

Other variations on the method have been reported. Mertz suggests a sinusoid fitting approach where the period of the carrier fringes is set to 3 pixels [66]. Similar algorithms to those used in the FFT method have also been presented by Womack [67].

The analysis of fringe data by both the phase stepping and Fourier transform approaches yields a wrapped phase map. The automatic processing of this image to produce a contiguous phase distribution is described in the following section.

1.3.6 Automatic Phase Unwrapping

Automatic phase unwrapping is a fundamental requirement for the industrial application of the phase stepping and Fourier transform methods. This field has seen a large amount of research activity in the last few years owing to the number of applications which require phase unwrapping as part of the analysis procedure. The task is to remove the phase wrapover points (or fringe edges) by adding or subtracting 2π from the pixels following the edge. In this way a contiguous phase distribution over the whole image can be produced. The problems faced by a phase unwrapping algorithm are basically due to the presence of bad data which may take many forms : random noise, object discontinuities, poorly modulated pixels, and insufficient resolution in the imaging system or frame capture device. A further requirement is for algorithms to produce meaningful information despite the presence of bad data and discontinuities.

The most well known algorithm for phase unwrapping was reported Macy [65] and operates along a raster of wrapped phase values. The wrapped phase at a point is given by $\phi(x,y)$, with the final unwrapped phase $\phi'(x,y)$ being defined as

$$\phi'(x,y) = \phi(x,y) + 2n\pi, \tag{1.20}$$

where n is an integer. The algorithm is based on the Nyquist sampling theorem which requires that one fringe be sampled over at least two pixels [30, p.67-68]. Therefore :

$$-\pi \leq \phi'(x_j,y) - \phi'(x_{j-1},y) \leq \pi. \tag{1.21}$$

Adding π to equation 1.21 and substituting equation 1.20 gives

$$0 \leq \phi(x_j,y) - \phi'(x_{j-1},y) + (2n+1)\pi \leq 2\pi. \tag{1.22}$$

This leads to a recurrence relationship for the unwrapped phase

$$\phi'(x_j,y) = AMOD\,[\phi(x_j,y) - \phi'(x_{j-1},y) + 201\pi, 2\pi] + \phi'(x_{j-1},y) - \pi. \tag{1.23}$$

The function AMOD produces the remainder of the first argument divided by the second, and the number 201 serves to keep the first argument positive for the correct functioning of the statement in Fortran. A similar procedure is then followed to determine the phase changes from one raster to another so that the two dimensional unwrapped phase can be determined.

The procedure works correctly providing that the Nyquist sampling criteria is satisfied over all points (equation 1.21). However, the presence of noise in the wrapped phase can lead to violations of this equation. If the phase at a point is in error then the phase may jump by 2π from one point to the next, jumping back again at the next pixel. Providing both discontinuities are located, the rest of the phase map remains unaffected by the noise. If the phase jumps by 2π but then does not return to the correct value, only one discontinuity (wrapover point) will be found and the unwrapped phase on the rest of the row will be in error. The probability of the phase not returning to its initial value after the noise point is increased when the slope of the phase is high. This argument was first presented by Creath [33]. The problem was solved in this case by applying filters to both the original cosinusoidal fringe fields and the wrapped phase map. This approach may suffice in certain cases but cannot be guaranteed for all fringe fields. The other problems of discontinuities in the object, such as in figure 1.16 must also be accounted for. Several manual methods, such as windowing, may be used to define object boundaries [68]. However, it is difficult to see how phase unwrapping along rows and then columns would accommodate the discontinuity in the fringe field of figure 1.16.

This discussion implies that a more complex phase unwrapping algorithm is required to successfully accommodate the problems of a general wrapped fringe field. The initial concept of unwrapping along rows followed by columns introduces streaks in the final phase map [33]. If the algorithm is altered to unwrap along the columns before the rows then the streaks will still be formed, but at right angles to the previous case. This can be demonstrated from the analysis by Itoh in one dimension [69], which was extended to the two dimensional case by Ghiglia [36]. The method is similar to that by Macy as it relies upon the sampling theorem but is represented in a different

manner. The one-dimensional case will be considered first. Let the unwrapped phase values be represented by $\phi(n)$ where n is an integer defining the position in the row of data. A wrapping operator W may then be defined as

$$W_l[\phi(n)] = \phi_{pw}(n) \tag{1.24}$$

where $\phi_{pw}(n)$ are principal values of the phase given in the wrapped map, and l is a label. This equation may be re-expressed as

$$W_l[\phi(n)] = \phi(n) + 2\pi k_l(n) \tag{1.25}$$

with k_l being a sequence of integers defined such that

$$-\pi \leq W_l[\phi(n)] \leq \pi. \tag{1.26}$$

A second operator is needed to define the difference in phase between two points :

$$\Delta\phi(n) = \phi(n) - \phi(n-1). \tag{1.27}$$

Applying the difference operator to the wrapping operator, equation 1.25, gives

$$\Delta W_1[\phi(n)] = \Delta\phi(n) + 2\pi\Delta k_1(n). \tag{1.28}$$

If this equation is then wrapped, we obtain the wrapped difference of wrapped phases :

$$W_2\{\Delta W_1[\phi(n)]\} = \Delta\phi(n) + 2\pi[\Delta k_1(n) + k_2(n)]. \tag{1.29}$$

The right hand side of this equation is a wrapped phase and therefore must have a value between $-\pi$ and π. Applying the Nyquist criteria to the phase difference $\Delta\phi(n)$

$$-\pi \leq \Delta\phi(n) \leq \pi, \tag{1.30}$$

the term $2\pi\left[\Delta k_1(n) + k_2(n)\right]$ is then equal to zero. Therefore

$$\Delta\phi(n) = W_2\left\{\Delta W_1\left[\phi(n)\right]\right\}, \tag{1.31}$$

which can be re-expressed as

$$\phi(m) = \phi(0) + \sum_{n=1}^{m} W_2\left\{\Delta W_1\left[\phi(n)\right]\right\}. \tag{1.32}$$

Equation 1.32 implies that the phase may be unwrapped in one dimension by integrating the wrapped differences of wrapped phases providing the Nyquist sampling theorem is not violated. If the Nyquist criteria is broken then the true phase cannot be recovered, but the unwrapped phase produced by the algorithm will have had any phase jumps greater than π removed.

The analysis may be extended into two dimensions (m, n) by using a two dimensional wrapping operator and two expressions for the phase difference in each of the two dimensions [36].

$$W_1\left[\phi(m,n)\right] = \phi(m,n) + 2\pi k_1(m,n), \tag{1.33}$$

$$\Delta_m\phi(m,n) = \phi(m,n) - \phi(m-1,n), \tag{1.34}$$

$$\Delta_n\phi(m,n) = \phi(m,n) - \phi(m,n-1). \tag{1.35}$$

Now two equations for the wrapped differences of wrapped phases can be formed in the same way as for equation 1.29 :

$$W_2\left\{\Delta_n W_1\left[\phi(m,n)\right]\right\} = \Delta_n\phi(m,n) + 2\pi\left[\Delta_n k_1(m,n) + k_2(m,n)\right], \tag{1.36}$$

$$W_3\left\{\Delta_m W_1\left[\phi(m,n)\right]\right\} = \Delta_m\phi(m,n) + 2\pi\left[\Delta_m k_1(m,n) + k_3(m,n)\right] \tag{1.37}$$

The Nyquist criteria must now be satisfied in two directions such that the second term on the right hand side of equations 1.36 1.37 both become zero.

$$-\pi \leq \Delta_n\phi(m,n) \leq \pi, \tag{1.38}$$

$$-\pi \leq \Delta_m \phi(m,n) \leq \pi. \tag{1.39}$$

Then

$$\Delta_n \phi(m,n) = W_2 \{\Delta_n W_1 [\phi(m,n)]\}, \tag{1.40}$$

$$\Delta_m \phi(m,n) = W_3 \{\Delta_m W_1 [\phi(m,n)]\}. \tag{1.41}$$

From these two equations it can be seen that any number of paths exist for unwrapping the phase between two points by following a sequence of sections along rows and columns. If the Nyquist criteria is satisfied for all points in an image, then the same phase difference between the two points will result independent of the path taken. However, if the Nyquist criteria is not fulfilled then the calculated phase difference will become path dependent. This situation may also arise owing to the presence of noise, as described above [33], and streaks will form according to the direction of unwrapping and the presence of bad data. The effect may be seen in figure 1.19 which is the unwrapped phase of figure 1.16 formed by unwrapping the left most column and then all the rows. The unwrapped phase values have been encoded to give a grey scale image where black indicates the minimum and white the maximum phase values. The underlying trend in the unwrapped phase can just be seen in this example but the presence of streaks makes it very difficult to achieve a proper quantitative interpretation. Path dependent unwrapping is undesirable as there should be a single result for the unwrapped phase map regardless of the path which is taken to form it. Therefore, the task of phase unwrapping in two dimensions requires a path independent method.

The first approach to consider path independent phase unwrapping in two dimensions was by Ghiglia et al [36] using cellular automata. The cellular automata inherently possess path independency due to the rule making of the automata which is not limited to one dimension, but considers all the neighbouring points to a pixel. The solution is attained by an iterative process. For each iteration the phase at every pixel may be changed according to a set of rules. In Ghiglia's work [36] the automata

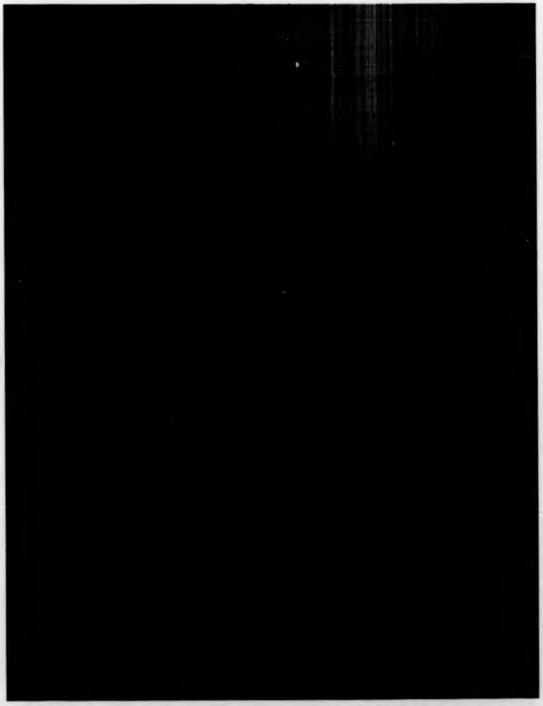


Figure 1.19: Unwrapped phase of Holographic Fringes using a Macy type Algorithm

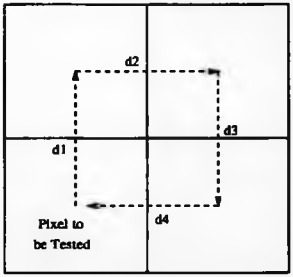
possessed an interaction length of one pixel, hence the four nearest neighbours to the point are considered. The phase differences are computed for each of these points with respect to the point in question, and then rounded to the nearest integer multiple of 2π. The sum of the integers is taken and the sign used to change the phase in the appropriate direction by 2π. This process forms one local iteration. In this way, the automata slowly moves toward a solution with the simple rules for individual points giving an overall complex collective behaviour. The automata will converge to an oscillatory state of period two iterations. The integer multiples of 2π are then averaged; this is called a global iteration. Further local iterations will produce an oscillatory state and further global iterations may then produce a consistent unwrapped solution.

The natural path independency of the automata is able to process around identified regions of the fringe field which should not be included in the analysis. This is achieved by excluding the identified points from the decision-making process of the automata for the valid pixels. For an irregularly shaped or discontinuous object (see figure 1.16 for example), the points outside the field may be defined by either a manual boundary definition or an automatic detection of the sinusoid modulation at each point (see section 1.1.2). The automata processes around these flagged points and therefore does not produce the streaks which Macy's and similar algorithms form.

In a similar way, if noise points or other data inconsistencies can be detected, then isolated areas of bad data can also be excluded from the unwrapping process of the automata. Ghiglia [36] proposes a simple data validity test for a 2×2 square of pixels, see figure 1.20. The phase is unwrapped along the closed path between the four points. If the result is zero then the four data points are consistent. Otherwise all four points are flagged as bad data and excluded from the decision making of the automata. This process has been applied to the wrapped map of figure 1.16 and the result is shown in figure 1.21, with the inconsistent regions highlighted in white, and the grey scale range of the wrapped phase reduced to 64 grey levels. It can be seen from the attempt at unwrapping using the first column and then each row, (figure 1.19), that these flagged inconsistencies give rise to streaks in the unwrapped

phase. The validity test functions well for detecting path inconsistencies produced by noise which often affect a single pixel.

2 * 2 Pixel Data Consistency Test



---------- Denotes Unwrapping Path

d1,...,d4 - Phase Differences

Sum (d1 + d2 + d3 + d4) = 0

if path is consistent

Figure 1.20: Valid Pixel Test for Phase Unwrapping

Other types of discontinuity can be caused by aliasing. In this case the inconsistency is part of a regional behaviour and is not confined to a 2×2 set of points. Therefore the path check over 4 pixels is not sufficient to detect the presence of large, connected inconsistent paths. The phase unwrapping process cannot converge as the addition of multiples of 2π cannot produce a smooth unwrapped phase distribution everywhere in the image. The automata enters a very long period cycle forming similar patterns but no two are identical. Ghiglia suggests a manual boundary definition

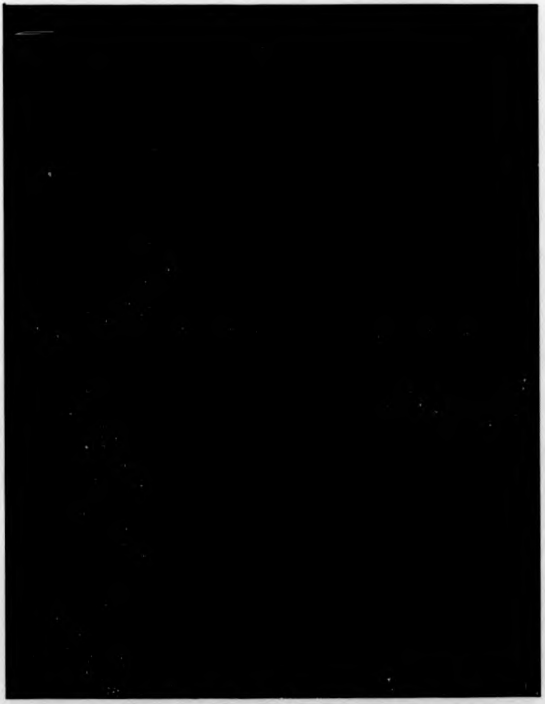


Figure 1.21: Holographic Wrapped Phase Map with Inconsistent Regions Highlighted in White

to force the automata to exclude more inconsistent paths and then converge to a solution.

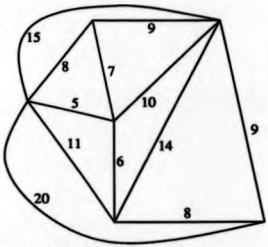
This work represents a substantial improvement in phase unwrapping methodology with many images becoming amenable to automatic analysis using cellular automata. The disadvantage of the method is due to the many iterations, both local and global, which are required for convergence. Therefore the execution time is several hours. This problem has recently been removed by Buckberry et al [70] by using a frame store with built in arithmetic processing. Other enhancements to the local neighbourhood automata rules were also made to increase the convergence rate [71]. The phase may be unwrapped over a 512×512 pixel image in less than 40 seconds using this system. A second, and more serious, disadvantage is the inability of the method to converge when aliasing type discontinuities are present. It must be stated, however, that a path independent solution for the unwrapped phase does not exist in this case. In both the work by Ghiglia [36] and Buckberry [70, 71] a manual region definition is needed to prevent certain paths being used for unwrapping.

The idea of path independent unwrapping has been pursued by Huntley [37]. The algorithm proposed also uses a 2×2 pixel test for data consistency, but in this case the sign of the *error* in the unwrapped phase is retained. Each negative inconsistency is then paired to the nearest positive inconsistency and a cut line defined between them. When all inconsistent data points have been paired the phase is unwrapped along any path that excludes the cut lines. As the points disregarded by both Ghiglia's and Huntley's algorithms are the same the two methods would be expected to produce similar results. The main difference is that Huntley's method eliminates all paths between neighbouring discontinuities and not just the discontinuities themselves. This may be an advantage in one of the examples given by Ghiglia where aliasing errors are present in the wrapped phase map. The points detected by the 2×2 pixel test form a ring [36, p.277]. If the sign of the bad data points oscillates around the ring then the cut lines would automatically define a circular boundary similar to that specified manually by Ghiglia [36, figure 12, p.278].

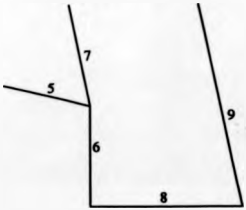
Another new method which produces nearly path independent results utilises a

minimum weight spanning tree (MST) construction for the unwrapping path. This method has been independently developed by both Ettemeyer [72] and Towers, Judge and Bryanston-Cross [73]. In this approach, the pixels are considered as vertices in a graph with edges linking neighbouring pixels into a grid. Each edge is assigned a weighting which may be defined as the local phase difference between the two vertices. A path (or tree) is then constructed which joins all vertices in such a way that the sum of the weights along the edges in the tree is a minimum. Therefore the pixels which are closest in phase to their neighbours are joined first, whilst pixels in regions of high phase gradient are added at the ends of the tree and cannot affect the rest of the phase map. The minimum weight spanning tree formed can accommodate discontinuities as in the automata method and is nearly path independent [74, p.61-64]. The path dependency can be examined with reference to figure 1.22. This shows a graph with six vertices and weighted edges linking the vertices. Two possible minimum weight spanning trees exist for this graph and are shown in the figure. Each MST has the same weight but is formed by a different path. This can only occur when there is a choice of more than one edge with minimum weight to be added to the tree. This phenomena is very unlikely when the weight computation is performed with real numbers as in the work by Towers et al [73].

A second significant advance in unwrapping strategies has been shown by using two scales to examine the fringe field. This was first implemented by Judge [75] using tiles. A tile is normally considered as a square region of pixels, typically 30 × 30. An unwrapping algorithm is initially applied to the wrapped phase data in each tile. The algorithm used can be based on the fringe counting method, employ cellular automata, or use the MST approach. This formulates a set of solutions for the individual regions. The tiles were initially joined according to the fit along neighbouring tile boundaries, and the starting tile found by trial and error. An MST approach was then adopted using the average fringe density in the tile and the edge fit as the weighting factors [76]. This gives a larger scale data consistency where the connection path is based on tile sized variations in fringe density and edge fit. Thereby the regional effects of aliasing errors can be accommodated at the tile level by appropriate selection



Original Graph Showing Vertices and Weighted Edges



Two Possible Minimum Weight Spanning Trees with Different Paths

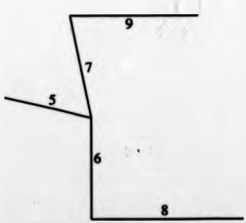


Figure 1.22: Examples of Minimum Weight Spanning Trees

of the tile size.

Two other methods for phase unwrapping have been described. Gierloff's method [38] considers dividing the fringe field into regions. The region boundaries are defined to be the wrapover points by an edge detection process resulting in a set of regions each containing contiguous data. The phase is then offset between whole regions rather than individual points. The calculations of the offsets are inherently more stable because of the number of points along the edge of a region than on a point-wise basis. However, the method has two disadvantages. Firstly, a large number of input parameters require specifying. Gierloff suggests that this allows tailoring of the algorithm to a particular fringe field to attain the correct result. This implies that the analysis cannot be carried out automatically and not by engineers unfamiliar with the software. Secondly, the algorithm depends on the correct detection of the wrapover points to assign a point to a region. Edge detection is defined by a user-supplied threshold and the number of neighbouring points which are in the region of interest. If a noise point is reached, then a region may be falsely extended beyond a wrapover edge. It may then be difficult to separate this new area into two regions because of the number of neighbouring points which are in the new area. This phenomena may explain the errors seen in the unwrapped phase maps presented in the paper.

Finally, Green et al [39], have proposed a phase unwrapping algorithm operating in the frequency domain in one dimension. In their method, it is assumed that the unwrapped phase has a limited bandwidth. If a wrapover point is inserted into the unwrapped phase then the bandwidth would be expected to increase. Conversely, the removal of a wrapover point from the wrapped phase would probably decrease the bandwidth. An unwrapping procedure was developed based on these ideas. The phase in the wrapped map is changed at each point by 2π. The change which results in the greatest reduction of bandwidth is selected and the process repeated. This approach is shown, both theoretically and by example, to fail in some cases as it is possible for the removal of a genuine wrapover point to increase the bandwidth in some cases. The work remains of interest owing to the possibility of combining frequency and spatially based unwrapping algorithms.

1.4 Summary of Analysis Procedures

It can be seen from the preceding discussion that the development of automatic fringe analysis systems requires the combination of optical, computing, and digital information processing technologies. The current state of hardware devices permits video rate frame capture to a resolution of $512 \times 512 \times 8$ bits. In practice this serves to examine many components in different industries : the automobile industry [8, 70, 77], in aerospace [7, 5], and in non-destructive testing [3]. Most of these applications consider the inspection of solid surfaces where the spatial frequency of the fringes encountered is sufficiently low so as not to violate the Nyquist sampling criteria. Also in each of these cases a quantitative understanding of the surface deformation is needed either for use in quality assessment, or comparison with theoretical predictions, or for experimental investigation.

From the methods and analysis procedures described above, the quasi-heterodyne or Fourier transform methods seem the most appropriate to provide full field information in an automatic manner. This conclusion is not evident from the literature which shows current research in all the analysis systems described. Apart from the quasi-heterodyne and Fourier transform methods, there is still considerable interest shown in fringe tracking and curve fitting methods to produce phase data at any point in the field. However, most of the recent work in this field proposes semi-automatic analysis systems (see for example [52, 51, 49]). This type of system is contradictory to the expected requirements of fringe analysis systems in the next decade as predicted by market research and presented in a paper by Jüptner [78]. The results of this research indicate that the majority of applications are in production line environments and that a quantitative analysis is required with the results going to non-optical specialists. This implies the need for fully automated data reduction. It is possible for the software analysis in fringe tracking and phase unwrapping to be fully automated. However, the result from phase unwrapping has two advantages over the fringe tracking approach. Firstly, phase unwrapping directly gives unambiguous phase measurements at a regular spaced grid of points without the need for further

processing and operator assignment of fringe orders. Secondly, in fringe tracking the phase values between fringe maxima or minima are found by interpolation, whereas the unwrapped phase gives measured values. This is a consequence of the phase stepping and FFT methods calculating the phase at every point in the field.

Therefore the quasi-heterodyne and Fourier transform methods leading to a wrapped phase map and subsequent phase unwrapping is the analysis sequence most likely to satisfy current and future industrial requirements. This advancement is achieved with an increased complexity in the optical system used.

The FFT approach has one disadvantage with respect to phase stepping owing to the requirement of obtaining sufficient carrier fringes to give a monotonic change in fringe order numbers. In the general case, this cannot be guaranteed for an unknown fringe field and the carrier frequency is often obtained by trial and error. The problem can be expressed equivalently as imposing a limit on the maximum phase gradient which can be correctly analysed in a carrier fringe image. The main advantage of the FFT method is that only a single image is required to perform the analysis. However, with the advances being made in computer power and storage capacity the importance of using a single image is being decreased. The majority of work in this thesis concerns the use and application of phase stepping methods as part of the analysis procedure. The relative merits of the Fourier transform and phase stepping methods will be discussed further during the course of the thesis.

The following chapter will present the necessary background to interferometry, holography and holographic interferometry.

Bibliography

[1] R. L. Powell, K. A. Stetson, "Interferometric vibration analysis by wavefront reconstruction", Journal of the Optical Society of America, volume 55, p.1593-1598, 1965.

[2] W. P. O. Jüptner, "Quantitative Evaluation of Industrial Components", Society of Photo-optical and Instrumentation Engineers, volume 746, p.3-7, 1987.

[3] C. Guedj, X. Briottet, "Non Destructive Testing by Electronic Speckle Interferometry", Society for Experimental Mechanics conference proceedings on Hologram Interferometry and Speckle Metrology, p.367-373, 1990.

[4] A. A. M. Maas, H. A. Vrooman, "In Plane Strain Measurement by digital Phase shifting Speckle Interferometry", Society of Photo-optical and Instrumentation Engineers, volume 1162, p.248-256, 1989.

[5] R. W. T. Preater, "In Plane Strain Measurement of High Speed Rotating Components using Double Pulsed Laser ESPI", Society for Experimental Mechanics conference proceedings on Hologram Interferometry and Speckle Metrology, p.288-293, 1990.

[6] A. Moore, J. Tyrer, "Surface Strain Measurement with ESPI Applied to Fracture Mechanics", Society for Experimental Mechanics conference proceedings on Hologram Interferometry and Speckle Metrology, p.192-198, 1990.

[7] R. J. Parker, D. G. Jones, "Industrial Holography the Rolls-Royce Experience", Society of Photo-optical and Instrumentation Engineers, volume 699, p.111-126,

1986.

[8] J. Davies, C. Buckberry, "Applications of a Fibre Optic TV Holography System to the Study of Large Automotive Structures", Society of Photo-optical and Instrumentation Engineers, volume 1162, San Diego 1989.

[9] K. Oh, R. J. Pryputniewicz, "Application of Electro-Optic Holography in the Study of Cantilever Plate Vibration with Concentrated Masses", Society for Experimental Mechanics conference proceedings on Hologram Interferometry and Speckle Metrology, p.245-253, 1990.

[10] B. P. Hildebrand, K. A. Haines, "Multiple Wavelength and Multiple Source Holography Applied to Contour Generation", Journal of the Optical Society of America, volume 57, number 2, p.155-162, 1967.

[11] P. Hariharan, B. F. Oreb, "Two Index Holographic Contouring : Application of Digital Techniques", Optics Communications, volume 51, number 3, p.142-144, 1984.

[12] M. Takeda, "Spatial Carrier Heterodyne Techniques for Precision Interferometry and Profilometry", Society of Photo-optical and Instrumentation Engineers, volume 1121, p.73-88, 1989.

[13] P. J. Bryanston-Cross, "High Speed Flow Visualisation", Progress in Aerospace Science, volume 23, number 2, p.85-104, 1986.

[14] B. Breuckmann, W. Thieme, "Computer Aided Analysis of Holographic Interferograms using the Phase Shift Method", Applied Optics, volume 24, number 14, p.2145-2149, 1985.

[15] P. J. Bryanston-Cross, J. W. Gardner, "Holographic Visualisation of a Combustion Flame", Optics and Lasers in Engineering, volume 9, p.85-100, 1988.

[16] B. L. Button, B. N. Dobbins, S. P. He, S. Kapasi, L. S. Wang, "Flame Temperature Measurement using ESPI and Holography", proceedings of Fringe Analysis

Special Interest Group (FASIG) meeting, p.321-322, September 1990.

[17] D. D. Lowenthal, A. L. Hoffman, "Quasi-Quadrature Interferometer for Plasma Density Radial Profile Measurements", Review of Scientific Instruments, volume 50, number 7, p.835-843, 1979.

[18] P. Hariharan, "Optical holography", Cambridge University Press, 1984.

[19] A. A. Michelson, "Studies in Optics", Phoenix books, University of Chicago Press, Third Impression 1962, chapters 1,2,3, and 7.

[20] C. M. Vest, "Holographic Interferometry", Wiley series in pure and applied optics, 1979.

[21] R. Dändliker, "Heterodyne Holographic Interferometry", in Progress in Optics, volume 17, ed. E. Wolf, p.1-84, North Holland, Amsterdam, 1980.

[22] J. H. Bruning, D. R. Herriott, J. E. Gallagher, D. P. Rosenfield, A. D. White, D. J. Brangaccio, "Digital Wavefront Measuring Interferometer for Testing Optical Surfaces and Lenses", Applied Optics, volume 13, number 11, p.2693-2703, 1974.

[23] P. Hariharan, "Quasi-heterodyne Hologram Interferometry", Optical Engineering, volume 24, number 4, p.632-638, 1985.

[24] J. C. Wyant, Applied Optics, volume 14, number 11, p.2622, 1975.

[25] D. Kerr, F. Mendoza-Santoyo, J. R. Tyrer, "Extraction of Phase Data from Electronic Speckle Pattern Interferometric Fringes using a Single Phase Step Method : a Novel Approach", Journal of the Optical Society of America, volume 7, number 5, p.820-826, 1990.

[26] P. Hariharan, B. F. Oreb, T. Eiju, "Digital Phase Shifting Interferometry : a Simple Error-Compensating Phase Calculation Algorithm", Applied Optics, volume 26, number 13, p.2504-2505, 1987.

[27] J. E. Grievenkamp, "Generalised Data Reduction for Heterodyne Interferometry", Optical Engineering, volume 23, number 4, p.350-352, 1984.

[28] R. Dändliker, R. Thalmann, "Heterodyne and Quasi-heterodyne Holographic Interferometry", Optical Engineering, volume 24, number 5, p.824-831, 1985.

[29] H. A. Vrooman, A. A. M. Maas, "New Image Processing Algorithms for the Analysis of Speckle Interference Patterns", Society of Photo-optical and Instrumentation Engineers, volume 1163, p.51-61, 1989.

[30] "Digital Signal Processing", ed. N. B. Jones, IEE Control Engineering Series, Volume 22, published by Peter Peregrinus Ltd., 1982.

[31] J. D. Trolinger, "Automated Data Reduction in Holographic Interferometry", Optical Engineering, volume 24, number 5, p.840-842, 1985.

[32] G. T. Reid, "Automatic Fringe Pattern Analysis : A Review", Optics and Lasers in Engineering, volume 7, p.37-68, 1986/7.

[33] K. Creath, "Phase Shifting Speckle Interferometry", Applied Optics, volume 24, number 18, p.3053-3058, 1985.

[34] J. C. Hunter, M. W. Collins, B. A. Tozer, "An Assessment of Some Image Enhancement Routines for use with an Automatic Fringe Tracking Programme", Society of Photo-optical and Instrumentation Engineers, volume 1163, p.83-94, 1989.

[35] J. Der Hovanesian, Y. Y. Hung, "Fringe analysis and Interpretation", Society of Photo-optical and Instrumentation Engineers, volume 1121, p.64-72, 1989.

[36] D. C. Ghiglia, G. A. Mastin, L. A. Romero, "Cellular-Automata method for Phase Unwrapping", Journal of the Optical Society of America, volume 4, number 1, p.267-280, 1987.

[37] J. M. Huntley, "Noise Immune Phase Unwrapping Algorithm", Applied Optics, volume 28, number 15, p.3268-3270, 1989.

[38] J. J. Gierloff, "Phase Unwrapping by Regions", Society of Photo-optical and Instrumentation Engineers, volume 818, p.2-9, 1987.

[39] R. J. Green, J. G. Walker, "Phase Unwrapping using a-priori Knowledge about the Band Limits of a Function", Society of Photo-optical and Instrumentation Engineers, volume 1010, p.36-43, 1988.

[40] Q. Yu, "Spin Filtering Processes and Automatic Extraction of Fringe Centrelines in Digital Interferometric Patterns", Applied Optics, volume 27, number 18, p.3782-3784, 1988.

[41] F. Becker, G. E. A. Meier, H. Wegner, "Automatic evaluation of interferograms", SPIE, volume 359, p.386-393, 1982.

[42] T. M. Kreis, W. P. O. Jüptner, "Fourier Transform evaluation of Interference Patterns : The Role of Filtering in the Spatial Frequency Domain", Society of Photo-optical and Instrumentation Engineers, volume 1162, 1989.

[43] Newport Information on ESPI System.

[44] T. M. Kreis, H. Kreitlow, "Quantitative analysis of holographic interferogram under image processing aspects", SPIE, volume 210, p. 196-202, 1979.

[45] B. L. Button, J. Cutts, B. N. Dobbins, C. J. Moxon, C. Wykes, "The identification of Fringe Positions in Speckle Patterns", Optics and Laser Technology, p.189-192, August 1985.

[46] T. Yatagai, S. Nakadate, M. Idesawa, H. Saito, "Automatic Fringe Analysis using Digital Image processing Techniques", Optical Engineering, volume 21, p.432-435, 1982.

[47] W. R. J. Funnell, "Image Processing Applied to the Interactive Analysis of Interferometric Fringes", Applied Optics, volume 20, number 18, p.3245-3250, 1981.

[48] D. W. Robinson, "Automatic Fringe Analysis with a Computer Image-Processing System", Applied Optics, volume 22, number 14, p.2169-2176, 1983.

[49] J. C. Hunter, M. W. Collins, B. A. Tozer, "A Scheme for the Analysis of Infinite Fringe Systems", Society of Photo-optical and Instrumentation Engineers, volume 1163, p.206-219, 1989.

[50] M. J. Matczak, J. Budzinski, "A Software System for Skeletonisation of Interference Fringes", Society of Photo-optical and Instrumentation Engineers, volume 1121, p.136-141, 1989.

[51] K. Liu, J. Y. Yang, "New Method of Extracting Fringe Curves from Images", Society of Photo-optical and Instrumentation Engineers, volume 1163, p.71-76, 1989.

[52] V. Parthiban, R. S. Sirohi, "Use of Grey Scale Coding in Labeling Closed Loop Fringe Patterns", Society of Photo-optical and Instrumentation Engineers, volume 1163, p.77-82, 1989.

[53] U. Mieth, W. Osten, "Three Methods for the Interpolation of Phase Values between Fringe Pattern Skeletons", Society of Photo-optical and Instrumentation Engineers, volume 1121, p.151-154, 1989.

[54] A. Sokolowski, "Automatic Determination of Regular Grid Distribution of Fringe Order in Interference Images", Society of Photo-optical and Instrumentation Engineers, volume 1121, p.159-161, 1989.

[55] N. A. Massie, "Real Time Digital Heterodyne Interferometry", Applied Optics, volume 19, p.154-160, 1980.

[56] C. Wykes, in Progress in Optics, ed. E. Wolf, volume 25, p.273-359, 1988

[57] P. V. Farrell, G. S. Springer, C. M. Vest, "Heterodyne Holographic Interferometry : Concentration and Temperature Measurements in Gas Mixtures", Applied Optics, volume 21, number 9, p.1624-1627, 1982.

[58] K. Creath, "Phase Measurement Interferometry Techniques", in Progress in Optics, ed. E. Wolf, volume 26, p.349-393, Amsterdam, 1988.

[59] M. Takeda, H. Ina, S. Kobayashi, "Fourier Transform Method of Fringe Pattern Analysis fro Computer based Topography and Interferometry", Journal of the Optical Society of America, volume 72, number 1, p.156-159, 1982.

[60] T. Kreis, "Digital Holographic Interference Phase Measurement using the Fourier Transform Method", Journal of the Optical Society of America A, volume 3, number 6, p.847-855, 1986.

[61] G. Frankowski, I. Stobbe, W. Tischer, F. Schillke, "Investigation of Surface Shapes Using Carrier Frequency Based Analysing System", Society of Photo-optical and Instrumentation Engineers, volume 1121, p.89-100, 1989.

[62] M. Kujawinska, J. Wojciak, "High Accuracy Fourier Transform Fringe Pattern Analysis", proceedings of Fringe Analysis Special Interest Group (FASIG) meeting, p.257-258, September 1990.

[63] D. R. Burton, M. J. Lalor, "Managing Some of the Problems of Fourier Fringe Analysis", Society of Photo-optical and Instrumentation Engineers, volume 1163, August 1989.

[64] T. R. Judge, C. Quan, P. J. Bryanston-Cross, "Holographic Deformation Measurements by Fourier Transform Technique with Automatic Phase Unwrapping", submitted to Optical Engineering.

[65] W. M. Macy, "Two Dimensional Fringe Pattern Analysis", Applied Optics, volume 22, number 23, p.3898-3901, 1983.

[66] L. Mertz, "Real Time Fringe Pattern Analysis", Applied Optics, volume 22, number 10, p.1535-1539, 1983.

[67] K. H. Womack, "Interferometric Phase Measurement using Spatial Synchronous Detection", Optical Engineering, volume 23, number 4, p.391-395, 1984.

[68] D. R. Matthys, J. A. Gilbert, T. D. Dudderar, K. W. Koenig, "A Windowing Technique for the Automated Analysis of Holo-Interferograms", Optics and Lasers in Engineering, volume 8, p.123-136, 1988.

[69] K. Itoh, "Analysis of the Phase Unwrapping Algorithm", Applied Optics, volume 21, number 14, p.2470, 1982.

[70] C. Buckberry, J. Davies, "The Application of TV Holography to Engineering Problems in the Automotive Industry", Society for Experimental Mechanics conference proceedings on Hologram Interferometry and Speckle Metrology, p.268-278, 1990.

[71] C. Buckberry - Private Communication.

[72] A. Ettemeyer, U. Neupert, H. Rottenkolber, C. Winter, "Schnelle und robuste Bildanalyse von Streifenmustern - ein wichtiger Schritt der Automation von holografischen Pruefprozessen", in Fringe '89 - Automatic Processing of Fringe Patterns, ed. W. Osten, R. J. Pryputniewicz, G. T. Reid, H. Rottenkolber, p.23-31, 1989.

[73] D. P. Towers, T. R. Judge, P. J. Bryanston-Cross, "Automatic Interferogram Analysis Techniques Applied to Quasi-Heterodyne Holography and ESPI", Optics and Lasers in Engineering, special issue 'Fringe Pattern Analysis', accepted for publication, 1991.

[74] N. Deo, "Graph Theory with Applications to Engineering and Computer Science", Prentice-Hall series in Automatic Computation, 1974.

[75] D. P. Towers, T. R. Judge, P. J. Bryanston-Cross, "Vibration Measurements using Dual Reference Beam Holography", Society of Photo-optical and Instrumentation Engineers, volume 1084, p.218-239, 1989.

[76] D. P. Towers, T. R. Judge, P. J. Bryanston-Cross, "A Quasi-Heterodyne Holographic Technique and Automatic Algorithms for Phase Unwrapping", Society of Photo-optical and Instrumentation Engineers, volume 1163, 1989.

[77] G. M. Brown, J. L. Sullivan, "The Computer Aided Holo-Photoelastic Method : Theory and Experiment", Society for Experimental Mechanics conference proceedings on Hologram Interferometry and Speckle Metrology, p.102-109, 1990.

[78] W. P. O. Jüptner, "Holographic Interferometry in the Next Decade", Society for Experimental Mechanics conference proceedings on Hologram Interferometry and Speckle Metrology, p.541-547, 1990.

Chapter 2

Interferometry, Holography and Holographic Interferometry

2.1 Interferometry

Optical interference effects are characterised by the superposition of light waves. If two light waves are considered, then the resultant intensity can be greater or less than the individual components depending on their relative phase. If both waves have the same amplitude and phase then the resultant amplitude will be doubled and the intensity quadrupled. Conversely, if the waves are of the same amplitude but opposite phases then the resultant amplitude is zero [1, chapter 2]. Interference effects are essentially the result of the superposition, or phasor sum, of the component waves.

Most practical interferometers consider the superposition of two light waves. The two waves are derived from the same light source in order to guarantee a constant phase relationship between the two-beams (see the following section 2.1.1). Several common characteristics are evident for all two beam interferometers [2]:

i) some means of splitting a light wave into two separate parts,

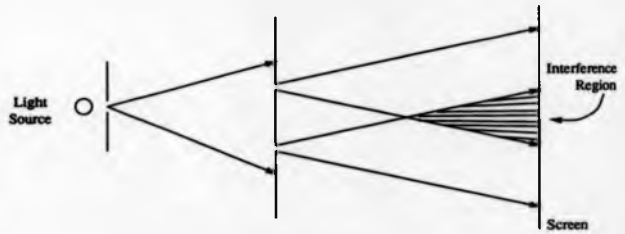
ii) the two waves follow different paths and acquire different phase lags,

iii) the two waves are recombined to observe interference fringes.

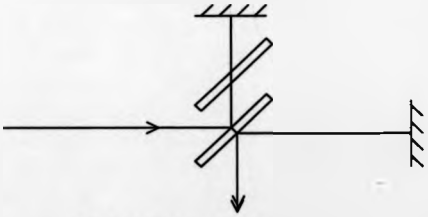
The first time recognition of the optical interference phenomena was in the formation of Newton's rings. This occurs when a thin film is illuminated, the interference occurring between the ray directly reflected from the film's outer surface, and the ray transmitted by the outer surface and reflected from the inner surface of the film. In this case, the two waves were formed from a single ray impinging the film surface. It was several years later that the two interfering waves were separated in the famous double slit experiment by Young, see figure 2.1 (a). Here the spatial extent of the source is limited by an aperture causing diffraction to a second screen containing two apertures a short distance apart. The resultant intensity distribution is observed on a screen some distance away. Interference fringes are seen because of the path difference between the two apertures and the point on the observation screen. This experiment proved the wave theory of light, as the two interfering waves could be acted upon independently and the result observed [1, chapter 2].

The configuration of subsequent interferometers was such that one of the interfering waves was modified by the object to be tested with the other wave used as a reference. The most relevant examples are the Michelson (then subsequently the Twyman-Green), and the Mach-Zehnder interferometers, see figure 2.1 (b) and (c) [3, p.224-234]. In each case a semi- transparent mirror acts as a beamsplitter and two independent wave paths are formed.

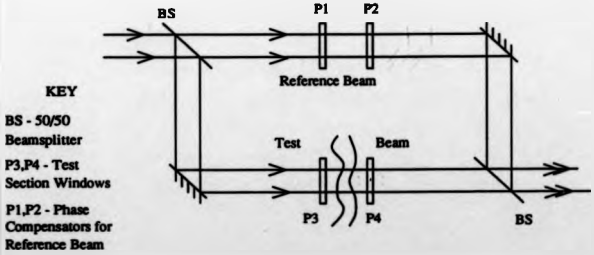
The applications of interferometers include : optical component testing (plane mirrors, aspherics, lens, etc.), and the examination of phase objects. Fringes are observed at the output of the interferometer due to the different phase lags experienced by each beam. However, the phase lags are integrated over the entire beam path. Hence perfect optical components are required in the reference beam path such that the fringes seen represent only the object inserted into the test beam. If this is not the case, a complex fringe pattern will result in which the components due to imperfections of the optics and the object to be measured cannot be separated. This was demonstrated by Vest using a Mach-Zehnder interferometer to examine a phase object [4, p.270] and demonstrates the major limitation of interferometers. A second problem is that only optically smooth (or specular reflecting) surfaces can be



(a) - Young's Double Slit Experiment



(b) - Michelson Interferometer



(c) - Mach-Zehnder Interferometer

Figure 2.1: Interferometers due to Young, Michelson, and Mach-Zehnder

examined with interferometers. Both these limitations were removed by the invention of holographic interferometry.

2.1.1 Light Sources for Interferometry and Holography

The idea of coherence is fundamental to the understanding of interference phenomena. For patterns of light and dark interference fringes to be observed the relative phases of the component waves must be constant over the observation time. This is the condition of temporal coherence. If the phase of a component light wave varies rapidly then the interference effects will average out and no fringes would be seen. A second condition, that of spatial coherence, is also required. This effect considers the transverse width of the illumination source. If the source width is large then light from different parts of the source will also give rise to unwanted interference effects and eventually the fringes will average out.

In practice, the advent of the laser with appropriate mode shape control provides a light source of good spatial and temporal coherence. The two parameters are interrelated. The output of a laser to be used for interferometry should oscillate in a single transverse mode. This is preferably the TEM_{oo} (Transverse Electromagnetic Mode) resulting in a Gaussian intensity cross section. The longitudinal oscillation modes in a laser need to be limited in order to achieve single frequency output. This is often achieved in practice by an intracavity etalon which selects a single line from the multifrequency output of the laser [5, p.69-73].

The temporal coherence attained from a laser relies on a comparison between two beams derived from the same source. Temporal coherence does not normally exist between two individual sources (laser or otherwise) and hence has only limited application to interferometry or holography.

The lasers normally used for interferometry and holography are tabulated in table 2.1.1.

Laser	Wavelength(nm)	Typical Power(mW)	Coherence Length
Helium-Neon	633	50	$\leq$ 20 cm
Argon-Ion	458	200	$\sim$ 20 m
	477	400	
	488	1000	
	514	1400	
Ruby	694	1 J pulsed	1 m
Nd:YAG (frequency doubled)	532	300 mJ pulsed	0.5 m
Laser Diode	670 to 800	$\leq$ 20	$\sim$ 20 m

The Helium-Neon (He-Ne) and Argon-ion lasers are both gas lasers and have a continuous wave output. The ruby and Nd:YAG (neodymium doped yttrium aluminium garnet) are solid state lasers and are operated in Q-switched mode, giving a short duration, 30 nanosecond (ns) and 10 ns respectively, pulsed output. Nd:YAG lasers can also be operated continuously. Laser diodes are a recent addition to the list of laser light sources. These devices are extremely compact and possess excellent coherence properties. New diodes with different wavelengths and greater output power are becoming available which implies an increased use of these devices in the future.

Physically, the coherence length limits the path difference between separate beams of an interferometer, and the size of object which may be studied. The coherence length is limited in most cases by the number of modes which are allowed to oscillate in the cavity.

For the purposes of this thesis it will be assumed that the maximum path length difference in the interferometer is less than the coherence length. The effect of fringe visibility reduction with increasing path length difference will be ignored.

2.2 Holography

The discovery of holography by Gabor in 1948 [6] marked a turning point in the development of optical concepts. For the first time, the phase and amplitude of an electromagnetic wave could be recorded and then reconstructed, giving a two stage 'wavefront reconstruction' process [7]. The significance of this was not fully realised until the mid 1960s when the advent of the laser and work by Leith and Upatnieks [8, 9, 10] produced wavefront reconstruction of three dimensional, diffusely reflecting objects. The basic concepts required to understand holography will be discussed below.

A holographic recording of an object can be appreciated when the restrictions of conventional photography are understood. A photograph forms a two-dimensional image of a three-dimensional scene via a lens. In this case, the light scattered by a point on an object is collected by the lens to form a focussed point in the image plane. Photosensitive film or paper placed in the image plane darkens according to the energy absorbed per unit area. The energy is proportional to the light intensity I, defined by

$$I = 2\langle \bar{v}.v \rangle$$

where $\bar{v}$ is the electric field vector, and $\langle\rangle$ denotes the time average [11, pages 6–10]. The time average of the wave is considered, as the observation time is large compared with the temporal period of the light. For a single beam of monochromatic light, the intensity is equal to the square of the amplitude of the electric field $I = v^2$. Hence the darkened grains in a developed photograph are dependent on a function of the intensity of the incident light, which in turn is proportional to v^2. Therefore a photograph contains no information regarding the phase of a light wave. The phase is related to the optical path length and thereby the distance from the object to the film. Hence the loss of phase information is synonymous to the loss of depth information in a photograph. It follows that in order to record an object or scene fully, the phase of the scattered light must be recorded as well as the intensity, i.e. a

complete wavefront reconstruction is needed.

The inspiration for Gabor's concept of wavefront reconstruction came from Bragg's X-ray microscope. The systems developed by Bragg and Gabor both depend on the separation of the two stages of diffraction which jointly cause image formation.

Abbe had initially described image formation by a lens as being the result of a double diffraction process [11, pages 35– 42] [5, chapter 1]. Abbe considered that light from an object is collected by a lens and forms a Fraunhofer, or far field, diffraction pattern in the lens's back focal plane. The second diffraction process forms when the Fraunhofer pattern is considered as a source, producing the pattern in the image plane.

Bragg's first diffraction process results from the spectral pattern formed by illumination of a crystal with X-rays. This is used to form a mask, and the second stage occurs when the mask is illuminated with visible coherent light. The holes in the mask cause diffraction of the incident light giving a magnified image of the crystal structure. The problem with the 'X-ray microscope' arises from the incomplete recording of the X-rays resulting from the first diffraction process. A standard photographic record is made and hence the phase information of the diffracted X-rays is lost. Bragg was able to demonstrate the principle, by appropriate selection of the type of crystal cell to be examined in which the first diffraction pattern is guaranteed to be real and of the same sign. In this case no information is lost in the photographic recording process and images of the crystal structure are successfully made.

Gabor aimed to achieve a complete recording of amplitude and phase by using a second beam in the recording process. The principle can be understood by considering the interference, or superposition, of two coherent monochromatic waves. Let the waves $\vec{a}_1$, $\vec{a}_2$ be defined by

$$\vec{a}_1 = a_1 \Re \left[e^{-i\phi_1} e^{-i2\pi ft} \right]$$

$$\vec{a}_2 = a_2 \Re \left[e^{-i\phi_2} e^{-i2\pi ft} \right]$$

where f is the optical frequency, ϕ_i, $i = 1,2$ is the phase and $\Re[\,]$ denotes the real part. The expressions incorporate the time-dependent optical oscillation $e^{-i2\pi ft}$, and the complex amplitude which is time independent $ae^{-i\phi}$. A photosensitive medium will receive an exposure dependent on the intensity, I, of the sum $\bar{a_1} + \bar{a_2}$. This simplifies to

$$I = \bar{a_1}\bar{a_1}^* + \bar{a_2}\bar{a_2}^* + \bar{a_1}\bar{a_2}^* + \bar{a_1}^*\bar{a_2} \tag{2.1}$$

$$I = I_1 + I_2 + 2\sqrt{I_1 I_2}\cos(\phi_1 - \phi_2) \tag{2.2}$$

where * implies the complex conjugate and I_i, $i = 1,2$ is the intensity of $\bar{a_i}$ (i.e. $I_i = a_i a_i^*$). Hence this intensity pattern also records the relative phase between the two waves from the third term of equation 2.2. This expression also applies to all two-beam interference patterns as described in section 2.1 and is seen to be of cosinusoidal form.

In holography, the two waves normally consist of light scattered by the object, and a reference wave which is often plane or spherical. For Gabor's system, also called in-line holography, a single beam illuminates a photographic negative (the object to be recorded), see figure 2.2. The object wave is the fraction of the incident light which is scattered by the darkened regions of the negative, and the reference wave is the directly transmitted part of the beam. Gabor realised that if the developed photographic record is re-illuminated with the wave $\bar{a_1}$, one of the elements in the reconstruction is proportional to the wave $\bar{a_2}$ in both phase and amplitude. Assuming photographic development such that the amplitude transmittance of the record is proportional to the illuminating intensity, then $\bar{w}$, the reconstructed wave, from equation 2.1, is given by

$$\bar{w} \propto \bar{a_1}I = \bar{a_1}I_1 + \bar{a_1}I_2 + \bar{a_1}\bar{a_1}\bar{a_2}^* + I_1\bar{a_2} \tag{2.3}$$

The last term in this equation is equal to the original wave $\bar{a_2}$ multiplied by a constant factor. The third term approximates to the conjugate wave of $\bar{a_2}$ given that the phase of the wave $\bar{a_1}$ is constant across the hologram (the conjugate or twin image).

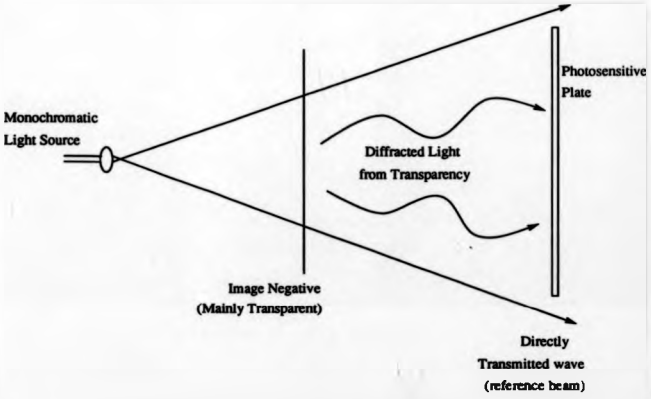


Figure 2.2: The Gabor system for In-Line Holography

In practice the Gabor approach had three main disadvantages :

i) the two image waves (and the other waves in equation 2.3) were superimposed in space,

ii) the development procedure required a positive hologram to be formed with accurately controlled exposure [11, pages 47–54],

iii) there was a limitation to using transparent objects with large clear areas to provide the reference wave.

However Gabor successfully demonstrated the three-dimensional nature of the reconstruction by making a hologram of a series of photographic negatives at different depths.

Interest in holography was re-opened by Leith and Upatnieks who eliminated the twin image problem [8, 9, 10]. The advent of the continuous wave (cw) laser at the same time was also of considerable experimental benefit because of the coherence produced, (see section 2.1.1). Their method used an application of carrier frequencies from communications theory. A separate, off-axis, reference beam was used and brought into the holographic plate at an angle, (see figure 2.3). The reference beam now has a spatially varying phase across the holographic plate. If the reference wave is plane and its angle to the plate is θ, then the beam may be described by

$$r(x,y) = re^{i2\pi\zeta_r x} \tag{2.4}$$

where $\zeta_r = \sin\theta/\lambda$ and the real part of the exponential is now assumed [5, pages 13– 16]. If equation 2.4 is substituted in equation 2.1 and $o(x,y) = o(x,y)e^{-i\phi(x,y)}$ is used to represent the object beam instead of $\bar{a_2}$ then

$$I_{off-axis} = r^2 + o(x,y)^2 + roe^{-i\phi(x,y)}e^{-i2\pi\zeta_r x} + roe^{i\phi(x,y)}e^{i2\pi\zeta_r x}. \tag{2.5}$$

From equation 2.5 it can be seen that the object wave is encoded as a phase and amplitude modulation of a set of carrier fringes given by ζ_r. Once again, if the

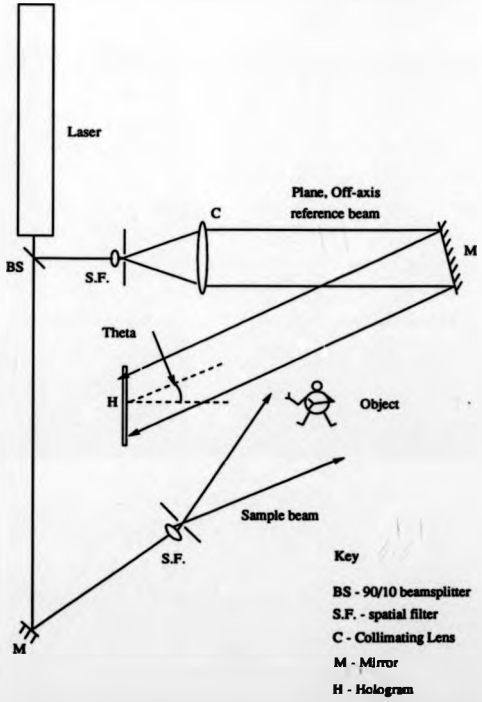


Figure 2.3: Holographic Recording System with Off-Axis Reference Beam

hologram is developed to give an amplitude transmittance proportional to the incident intensity and the image is reconstructed with $\bar{r}$, then $\bar{w}$, the reconstructed wave, is given by

$$\bar{w} \propto \bar{r} I_{off-axis} \tag{2.6}$$

$$\propto r^3 e^{i2\pi\zeta_c x} + o(x,y)^2 r e^{i2\pi\zeta_c x} + r^2 o(x,y) + r^2 \bar{o}^2(x,y) e^{i4\pi\zeta_c x}. \tag{2.7}$$

The first term in equation 2.7 is an attenuated reference beam with the second term forming a 'halo' around it. The third term reconstructs the original object wave, apart from a scalar factor. This wave is derived from the original object position and hence is virtual. The final term represents a conjugate, real image of the object. Both virtual and real images of the original object wave are spatially separated from each other, and from the directly transmitted wave and halo. This is apparent from the exponential factors in each of the reconstructed waves, see figure 2.4.

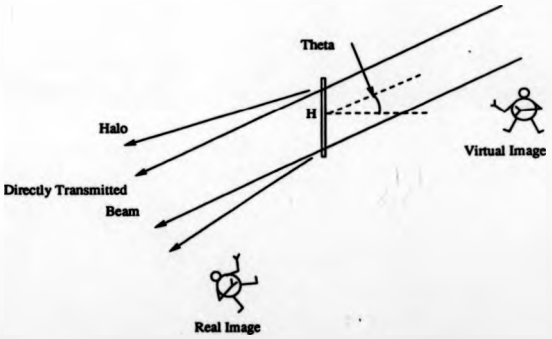


Figure 2.4: Reconstructed waves from an Off-Axis Hologram

Using off-axis holography, it became possible to examine both the virtual and

real images independently. This method has further advantages as the development procedure does not have to produce a positive hologram to obtain a positive image [11, pages 47–54].

No assumptions have been made regarding the form of the object, hence the method can be applied to both specularly reflecting and diffusely scattering surfaces.

In off-axis holography, the resolution required depends on the mean grating period formed by the off-axis reference beam. If the reference beam is incident at an angle θ to the holographic plate, with light of wavelength λ, then the period of the carrier fringes, T, is [4, p.46]

$$T = \frac{\lambda}{\sin\theta}. \tag{2.8}$$

Equivalently this can be written as a frequency, $f = \frac{1}{T}$,

$$f = \frac{\sin\theta}{\lambda}. \tag{2.9}$$

If $\theta = 30$ degrees and $\lambda = 633$ nm (i.e. a He-Ne laser), then $f = 790$ lines/mm. This represents a minimal requirement; to resolve higher carrier fringes and accommodate objects which scatter light over large angles a resolution of up to 2000 lines/mm is needed. Many emulsions have been developed for this purpose. However, their sensitivity is low compared with photographic film as a consequence of the high spatial resolution (see for example [4, p.47]).

2.3 Holographic Interferometry

The application of holography to industrial problems began concurrently in several parts of the world. Powell and Stetson's work is typical [12, 13]. The fundamental approach was to record more than one state of an object in the hologram. On reconstruction, each object state is formed in the image as though the original hologram is made up of a set of optically non-obtrusive surfaces. The reconstructed states of the object are mutually coherent and hence form interference fringes which map the change of state. The first holographic interferograms described by Powell and Stetson

were formed by the 'time average' approach [12]. A single timed exposure is made of a vibrating surface with a cw laser. If the object is assumed to vibrate sinusoidally, then the reconstructed fringe pattern takes the form of a zero-order Bessel function [12].

The development of the other techniques for holographic interferometry occurred rapidly. The major methods are double exposure, time-average, real time and stroboscopic interferometry [5, chapters 14,15]. Further advanced methods have also been developed for special cases. The different techniques give rise to different fringe functions which can be collectively described by :

$$\bar{w}_3 \propto r^2(\bar{o}(x,y)M_T(x,y)), \tag{2.10}$$

$$\propto \bar{o}(x,y)M_T(x,y), \tag{2.11}$$

where $M_T(x,y)$ is the characteristic function of the fringes, and $\bar{w}_3$ is the virtual image wave, from the third term of equation 2.7. In each of the following sections the characteristic function will be derived for the main interferometric techniques. A summary of the results is shown in figure 2.5.

2.3.1 Double Exposure Holographic Interferometry

In this case two separate holographic exposures are made of an object. Most of the work presented here concerns either static deformations or surface vibrations. In both cases two exposures are made. For static deformations the object's state is changed between the two exposures. The vibration analysis is achieved practically by using a double pulsed ruby laser. The vibrational case is a close approximation to the static case when a pulsed laser is used. This effect is due to the pulse duration, typically 30 ns, which freezes the object motion at a particular point in time.

Both these methods yield cosinusoidal fringe patterns as shown below [5, chapters 14,15]. The initial object wave in the hologram plane (x,y) is defined by

$$o_1 = o_1(x,y)e^{-i\phi_1(x,y)} \tag{2.12}$$

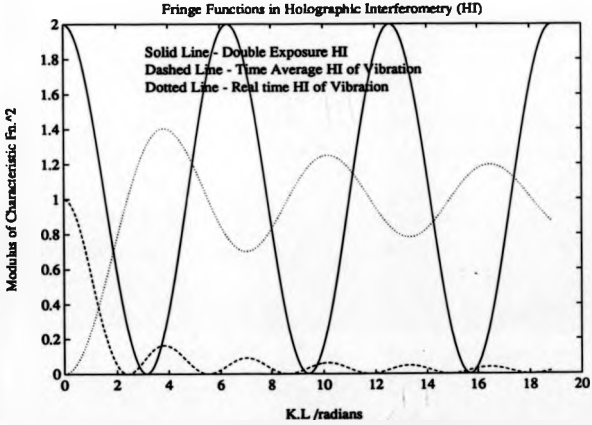


Figure 2.5: Fringe Functions for Different types of Holographic Interferometry

and the deformed object wave by

$$\bar{o}_2 = o_2(x,y)e^{-i\phi_2(x,y)}. \tag{2.13}$$

If the same reference beam $\bar{r}(x,y)$ is used for each exposure then the intensities forming the two holographic images are

$$I_1 = (\bar{o}_1(x,y) + \bar{r}(x,y))(\bar{o}_1(x,y) + \bar{r}(x,y))^* \tag{2.14}$$

and

$$I_2 = (\bar{o}_2(x,y) + \bar{r}(x,y))(\bar{o}_2(x,y) + \bar{r}(x,y))^* \tag{2.15}$$

The derivation given from equations 2.4– 2.7 may now be duplicated with $\bar{o} = \bar{o}_1 + \bar{o}_2$. Hence the wave representing the virtual image, given by the third term of equation 2.7, $\bar{w}_3$ may be expressed as :

$$\bar{w}_3 \propto r^2(\bar{o}_1 + \bar{o}_2) \tag{2.16}$$

This may be expanded using equations 2.12 and 2.13 to give

$$\bar{w}_3 \propto \left(o_1(x,y)e^{-i\phi_1(x,y)} + o_2(x,y)e^{-i\phi_2(x,y)}\right) \tag{2.17}$$

In practice the main effect of the surface translation is to change the phase of the object wave, hence $o_1(x,y) \simeq o_2(x,y) \simeq o(x,y)$ [5, 12, chapters 14,15], and

$$\bar{w}_3 \propto o(x,y)\left(e^{-i\phi_1(x,y)} + e^{-i\phi_2(x,y)}\right) \tag{2.18}$$

This defines the characteristic function, $M_T(x,y)$, for the double exposure case as :

$$M_T(x,y) = \left(1 + e^{-i(\phi_2(x,y) - \phi_1(x,y))}\right) \tag{2.19}$$

The observer will detect the intensity of the wave, I_{w_2}, given by

$$I_{w_2} \propto o(x,y)^2 \left(e^{-i\phi_1(x,y)} + e^{-i\phi_2(x,y)}\right)\left(e^{i\phi_1(x,y)} + e^{i\phi_2(x,y)}\right), \qquad (2.20)$$

$$I_{w_2} \propto 2o(x,y)^2 \left(1 + \cos(\phi_1 - \phi_2)\right). \qquad (2.21)$$

This expression illustrates the cosinusoidal structure of the reconstructed fringe pattern as discussed in the preceding sections, (see figure 2.5). From this expression, it can be seen that any phase distortion which is present in both exposures is cancelled out of the final fringe pattern. Therefore it is possible to use non-optical quality test section windows in holographic interferometry, (compare with section 2.1). Similarly, any distortion in chemically processing the hologram will apply to both images recorded and hence will not be seen in the fringe pattern. These characteristics make double exposure holographic interferometry the most widely used and simplest approach to obtaining quantitative information.

The form of the fringe pattern is the same for both specularly reflecting and diffusely reflecting surfaces. In the diffuse case, however, two effects are evident :

i) a high frequency intensity variation, or speckle pattern, due to the microstructure of the surface,

ii) a low frequency sinusoidal variation due to the deformation of the surface between the two exposures.

This can be seen from the following derivation which incorporates a speckled form of the interfering waves. Assuming that the hologram acts as a perfect wavefront storage component and the virtual image is viewed independently of the other reconstructed waves, then the interfering wavefronts can be considered as coming directly from the two object states. To account for speckle effects, the surface can be considered as a set of point scatterers [14, p.74-79]. The light reaching a point in the image plane is then the phasor sum of the light scattered from the number of points in the resolvable area of the object. In general this will consist of n scatterers or speckles. The light wave arriving at a point in the image plane from the undeformed object, u_1, can be

represented by

$$\bar{u_1} = u \sum_{p=1}^{n} e^{i\psi_p}, \tag{2.22}$$

where u is the amplitude of an individual point scatterer, and ψ_p is the phase of the pth speckle. After displacement the light arriving at the same point in the image plane is

$$\bar{u_2} = u \sum_{p=1}^{n} e^{i(\psi_p + \phi_p)}, \tag{2.23}$$

where ϕ_p is the phase change of the pth speckle caused by the surface deformation. The resulting intensity, I, is given by

$$\begin{aligned} I &= (\bar{u_1} + \bar{u_2}).(\bar{u_1}^* + \bar{u_2}^*), \\ &= u^2 \left\{ \sum_{p=1}^{n} \sum_{q=1}^{n} \left[e^{i(\psi_p - \psi_q)} + e^{i(\psi_p + \phi_p - \psi_q - \phi_q)} + \right. \right. \\ &\quad \left. \left. e^{i(\psi_p - \psi_q + \phi_p)} + e^{-i(\psi_p - \psi_q + \phi_p)} \right] \right\}, \\ &= u^2 \left\{ \sum_{p=1}^{n} \sum_{q=1}^{n} \left[e^{i(\psi_p - \psi_q)} + e^{i(\psi_p + \phi_p - \psi_q - \phi_q)} + 2\cos(\psi_p - \psi_q + \phi_p) \right] \right\} \end{aligned} \tag{2.24}$$

This expression illustrates the cosinusoidal fringe form which is modulated by random speckle phase. The mean value can be found by separating the summation for the terms where $p = q$ and $p \neq q$.

$$\begin{aligned} I &= u^2 \left\{ \sum_{p=1}^{n} \left[e^{i(\psi_p - \psi_p)} + e^{i(\psi_p + \phi_p - \psi_p - \phi_p)} + 2\cos\phi_p \right] \right\} + \\ &\quad u^2 \left\{ \sum_{p=1}^{n} \sum_{q \neq p} \left[e^{i(\psi_p - \psi_q)} + e^{i(\psi_p + \phi_p - \psi_q - \phi_q)} + 2\cos(\psi_p - \psi_q + \phi_p) \right] \right\} \end{aligned} \tag{2.25}$$

If n is large and the spatial average over a line of constant ϕ_p is taken, the terms $\sum_{p=1}^{n} \cos\psi$ and $\sum_{p=1}^{n} \sin\psi$ tend to zero when ψ varies randomly. Therefore

$$\langle I \rangle = 2u^2 n \left[1 + \cos\phi_p \right], \tag{2.26}$$

where $\langle\rangle$ denotes the spatial average. A cosinusoidal fringe pattern is observed given that the deformation phase ϕ_p is nearly constant over the image point. Even in the

case where the phase is not constant, fringes are still seen but of reduced contrast. This can be seen by substituting $\phi_p = \phi_o + \Delta\phi_p$ in equation 2.26 where ϕ_o is the centre value of ϕ_p.

$$\begin{aligned} \langle I \rangle &= 2u^2\langle\left[n + \sum_{p=1}^{n} \cos(\phi_o + \Delta\phi_p)\right]\rangle, \\ &= 2u^2\langle\left[n + \cos\phi_o \sum_{p=1}^{n} \cos\Delta\phi_p - \sin\phi_o \sum_{p=1}^{n} \sin\Delta\phi_p\right]\rangle. \end{aligned} \tag{2.27}$$

In the case where the phase variation $\Delta\phi_p$ is small, i.e. $\Delta\phi_p \ll \pi$, then the $\cos\Delta\phi_p$ terms are approximately equal to one, and the $\sin\Delta\phi_p$ are equal to zero. Hence

$$\langle I \rangle \approx 2u^2 n\,[1 + \cos\phi_o]\,. \tag{2.28}$$

Therefore the fringe pattern maps the variation of ϕ_o, the deformation phase at the central point of each resolved area. For good fringe contrast it is necessary for the speckle patterns to be essentially the same, differing only by a large scale systematic variation in phase, ϕ_o [4, p.57-64]. Hence the magnitude of deformation which can be measured using this technique is limited because of the requirement that the speckle patterns, and hence the surface microstructures, are correlated.

An example of a double exposure holographic interferogram is shown in figure 2.6, which shows one of the vibrational modes of a trumpet bell. The trumpet bell was sprayed to obtain a matt white finish in order to eliminate bright spots appearing in the image. This hologram was recorded with a double pulsed ruby laser with a pulse separation of 400 microseconds.

The object deformation can be related to the phase measured $\phi_1 - \phi_2$ from an analysis by Vest, see figure 2.7 [4, p.68-72]. Illumination is considered from O to a point P on the object which translates by $\underline{L}$ between the two exposures. Position Q is a viewing point on the hologram, and vectors $\underline{k}_1$ ..... $\underline{k}_4$ are propagation vectors of magnitude $\frac{2\pi}{\lambda}$. The phases of the two rays which reach Q from the points P and P1

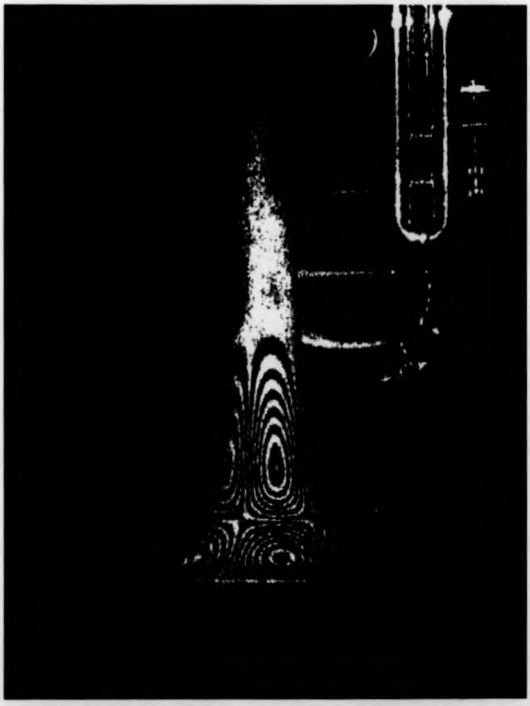
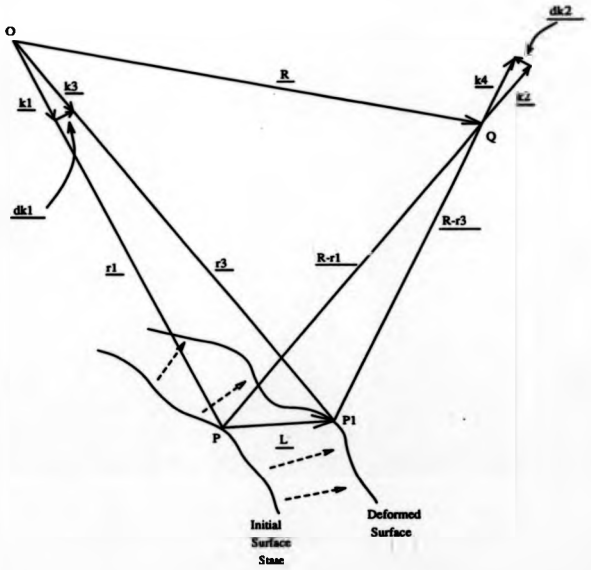


Figure 2.6: Double Pulsed Interferogram of a Vibrational Mode of a Trumpet Bell



Key

O - Illumination source

P - Initial position of a point on the object

P1 - Deformed position of point P

L - Vector translation of P

Q - Viewing Position from Hologram

Figure 2.7: Definitions to Determine the Sensitivity Vector

are given by :

$$\phi_1 = \underline{k}_1 \cdot \underline{r}_1 + \underline{k}_2 \cdot (\underline{R} - \underline{r}_1) + \phi_O, \tag{2.29}$$

$$\phi_2 = \underline{k}_3 \cdot \underline{r}_3 + \underline{k}_4 \cdot (\underline{R} - \underline{r}_3) + \phi_O, \tag{2.30}$$

where ϕ_O is the arbitrary phase of the ray at O. Hence the phase difference, δ, measured by an observer at Q is

$$\delta = \phi_2 - \phi_1. \tag{2.31}$$

By defining the small changes in the propagation vectors between the observation directions $\underline{k}_2, \underline{k}_4$, and the illumination directions $\underline{k}_1, \underline{k}_3$ as

$$\underline{k}_3 = \underline{k}_1 + \Delta\underline{k}_1, \tag{2.32}$$

$$\underline{k}_4 = \underline{k}_2 + \Delta\underline{k}_2, \tag{2.33}$$

the phase difference δ may be expressed as

$$\delta = (\underline{k}_1 - \underline{k}_2) \cdot (\underline{r}_3 - \underline{r}_1) + \Delta\underline{k}_1 \underline{r}_3 + \Delta\underline{k}_2 (\underline{R} - \underline{r}_3). \tag{2.34}$$

In most real systems the magnitude of $\underline{r}_1$ and $\underline{r}_3$ are much larger than the magnitude of $\underline{L}$. Therefore $\Delta\underline{k}_1 \perp \underline{r}_3$ and $\Delta\underline{k}_2 \perp (\underline{R} - \underline{r}_3)$ and the two last two terms on the right hand side of equation 2.34 are equal to zero. The phase difference measured from Q is then

$$\delta = (\underline{k}_1 - \underline{k}_2)\underline{L}, \tag{2.35}$$

$$= \underline{K} \cdot \underline{L}, \tag{2.36}$$

where $\underline{K} = \underline{k}_1 - \underline{k}_2$ is the sensitivity vector. This vector lies along the bisector of the illumination and observation directions. Therefore by appropriate choice of optical arrangement, the sensitivity of the phase measurements to different motions of the

surface can be achieved [4, p.150].

2.3.2 Time-Average Holographic Interferometry

This approach uses a single timed exposure from a cw laser to examine a vibrating surface. If the motion of the surface is assumed to be simple harmonic and consists of a single frequency ω then the time-dependent displacement of a point P(x,y) can be written as [5, p.232-234] :

$$\underline{L}(x,y,t) = \underline{L}(x,y)\sin(\omega t). \tag{2.37}$$

If the optical system is such that the sensitivity vector is $\underline{K}$, then the phase change $\Delta\phi(x,y,t)$ of light scattered from P is

$$\Delta\phi(x,y,t) = \underline{K}.\underline{L}(x,y)\sin(\omega t). \tag{2.38}$$

Let the light scattered from the mean position of the surface be described by

$$o(x,y) = o(x,y)e^{-i\phi(x,y)}, \tag{2.39}$$

then the scattered light from the vibrating surface is

$$\begin{aligned} \bar{o}(x,y,t) &= o(x,y)e^{-i(\phi(x,y)+\Delta\phi(x,y,t))}, & (2.40)\\ &= o(x,y)e^{-i(\phi(x,y)+\underline{K}.\underline{L}(x,y)\sin(\omega t))} & (2.41) \end{aligned}$$

Assuming linear processing of the hologram and reconstruction with a wave identical to the original reference wave, then the virtual image wave is proportional to the time average of $o(x,y,t)$ over the exposure time T.

$$\begin{aligned} \bar{w}_3 &\propto \frac{1}{T}\int_0^T o(x,y)e^{-i(\phi(x,y)+\underline{K}.\underline{L}(x,y)\sin(\omega t))}dt, & (2.42)\\ &\propto o(x,y)e^{-i\phi(x,y)}\frac{1}{T}\int_0^T e^{-i\underline{K}.\underline{L}(x,y)\sin(\omega t)}dt, & (2.43) \end{aligned}$$

$$\propto \quad \bar{o}(x,y)M_T(x,y), \tag{2.44}$$

with

$$M_T(x,y) = \frac{1}{T}\int_0^T e^{-i\underline{K}.\underline{L}(x,y)\sin(\omega t)}dt. \tag{2.45}$$

Normally the exposure time is set to be much greater than the temporal period of vibration, $T \gg \frac{2\pi}{\omega}$. Then

$$M_T(x,y) = \lim_{T\to\infty}\frac{1}{T}\int_0^T e^{-i\underline{K}.\underline{L}(x,y)\sin(\omega t)}dt, \tag{2.46}$$

$$= J_0[\underline{K}.\underline{L}(x,y)], \tag{2.47}$$

where J_0 is the zero order Bessel function of the first kind. The intensity which is observed in the virtual image is then

$$I_{\omega_2} \propto |\bar{o}(x,y)M_T(x,y)|^2, \tag{2.48}$$

$$\propto o(x,y)^2 J_0^2[\underline{K}.\underline{L}(x,y)]. \tag{2.49}$$

where $o(x,y)^2$ is the intensity of the object at rest. The image shows a bright fringe for the nodal line and fringes of decreasing intensity as the vibrational amplitude increases, (see figure 2.5). This method has a disadvantage in that the fringes lose contrast at high deformations (owing to the form of the Bessel function), which is partially offset by the brightest fringe lying along the nodal lines of the vibration.

An example of this type of fringe pattern is shown in figure 2.8 which shows one of the vibrational modes of a model cooling tower. The hologram was produced using a ruby laser operating in open lase mode, i.e. the Pockels cell was turned off. This has the effect of producing a series of pulses over a period of ~ 2 milliseconds which approximates to a time average exposure with a cw laser.

Figure 2.8: Time Average Interferogram of a Cooling Tower Model

2.3.3 Real Time Holographic Interferometry

Real time holographic interferometry refers to the case where one of the interfering wavefronts is generated by a hologram with the other formed live, or in real time [5, p.234-236]. A first exposure is normally made of the stationary object. The hologram is then developed and replaced in precisely the same position as for recording. The reconstructed virtual image wave is then

$$\tilde{w}_3 = -\tilde{o}(x,y) \tag{2.50}$$

where the minus sign is due to the hologram being a photographic negative [4, p.197-200] (this does not affect the fringes in the time average or double exposure methods). The intensity of the reconstruction beam is varied such that the virtual image wave from the hologram and the live wave from the object are equal in intensity. A fringe pattern will be observed when the object is deformed in some manner.

In the case where the deformation is static, the interference occurs between the reconstructed virtual image wave and the deformed object, given by $\tilde{o}(x,y).e^{i\delta(x,y)}$, or equivalently $\tilde{o}(x,y).e^{i\underline{K}.\underline{L}(x,y)}$. Therefore the intensity in the fringe pattern is

$$\begin{aligned} I(x,y) &\propto \left(-\tilde{o}(x,y) + \tilde{o}(x,y)e^{i\delta(x,y)}\right).\left(-\tilde{o}^*(x,y) + \tilde{o}^*(x,y)e^{-i\delta(x,y)}\right) & (2.51)\\ &\propto o(x,y)^2.\left(2 - e^{i\delta(x,y)} - e^{-i\delta(x,y)}\right), & (2.52)\\ &\propto o(x,y)^2.2\left\{1 - \cos\left(\underline{K}.\underline{L}(x,y)\right)\right\}. & (2.53) \end{aligned}$$

The fringe pattern is therefore of the same form as the double exposure case, but the fringe corresponding to zero motion is dark instead of bright, i.e. a dark field interferogram is produced.

If the object vibrates sinusoidally whilst being observed using the real time method, then the intensity seen will become time dependent;

$$I(x,y) \propto o(x,y)^2.2\left\{1 - \cos\left[\underline{K}.\underline{L}(x,y)\sin\left(\omega t\right)\right]\right\}. \tag{2.54}$$

Providing the temporal period of the vibration is much less than the observation time, T, (0.04 s for the human eye) then a time averaged fringe pattern results with

$$M_T^2 = \frac{1}{T}\int_0^T \{1-\cos[\underline{K}.\underline{L}(x,y)\sin(\omega t)]\}\,dt, \tag{2.55}$$

$$= 1 - J_0[\underline{K}.\underline{L}(x,y)]. \tag{2.56}$$

This is again a dark field interferogram with half the number of fringes as the time average case, (see figure 2.5).

Practically, the real time methods are difficult to implement. This is mainly owing to the problem of exact repositioning of the hologram. If the hologram is not replaced in precisely the same position for recording then extraneous fringes will be seen in the image. For quantitative measurements, it is desirable to remove these fringes and obtain a uniform background. This problem can be removed by using a liquid gate in which all chemical processing is performed in situ. A second difficulty is the effect of processing on the emulsion (assuming conventional silver halide photographic plates are used). During chemical processing the emulsion undergoes expansion and shrinkage. If the development is not uniform then a variable phase shift will be present in the reconstruction causing extraneous fringes. Both problems are removed when thermoplastic materials are used to record the hologram [15]. Development is performed in situ by the application of an electric field. In this case, there is no possibility of shrinkage effects in the equivalent of the emulsion layer.

2.3.4 Stroboscopic Holographic Interferometry

Stroboscopic illumination has traditionally been used to convert Bessel function (or time average fringes) into cosinusoidal fringes when a cw laser is used to examine a vibrating surface. The exposure from the laser is restricted to two instants of the vibration period, for example at the maximum and minimum displacements. The sequence of pulses are repeated until the hologram exposure attains the required level. If the exposure times are sufficiently short and the optical system is stable over

the exposure time, then the fringes observed in the reconstruction will be cosinusoidal and the method is equivalent to using a double pulsed laser. In practice the exposure times are finite which has the effect of reducing the fringe visibility [4, p.203-209].

The method can also be applied to the real time study of vibrations with a cw laser. An initial exposure is made of the static surface and the hologram processed and returned to its original position. The object is then stroboscopically illuminated and viewed through the reconstructed hologram.

2.3.5 Combining Quasi-Heterodyne Analysis with Holographic Interferometry

To apply quasi-heterodyne analysis some means of phase shifting the fringe pattern is required. In double exposure holographic interferometry this has been achieved by using dual reference beams to record and reconstruct the hologram, see figures 2.9 2.10.

Hologram recording is performed by using the first reference beam to record the initial state of the object and the second reference beam for the deformed state. For double pulsed vibration analysis a means is required to switch reference beams between the two exposures from a double pulsed laser. This is discussed in the applications chapter of the thesis. For reconstruction, both beams are used to form the two images of the object independently. By modulating the phase of one beam with respect to the other, the relative phase of the two beams is altered, and consequently the phase of the fringe field can be varied.

The use of dual reconstruction beams requires a stabilised optical bench as any environmental changes which affect only one of the reconstruction beams will alter the phase of the fringe field. This effect is minimised by computer control of image capture and phase shifting to reduce the time needed to acquire the phase stepped fringe fields.

The mechanisms available for phase modulation have been reviewed by Creath [16]and Wykes [2]. These can be summarised in the items below.

i) Moving an optical component such as a mirror or wedge, or rotation of

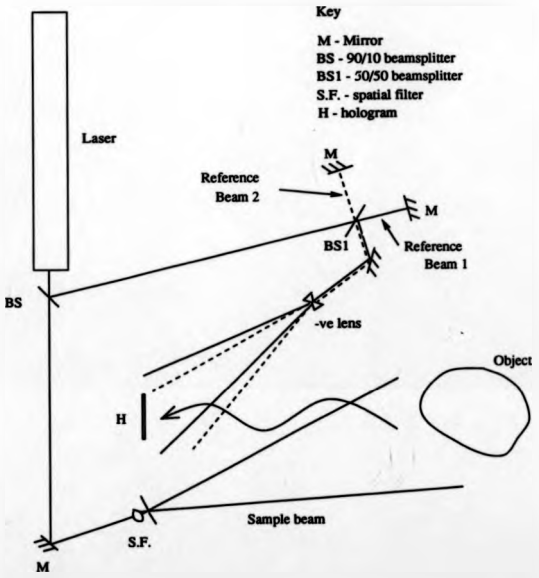


Figure 2.9: Recording System for Dual Reference Beam, Double Exposure Holographic Interferometry

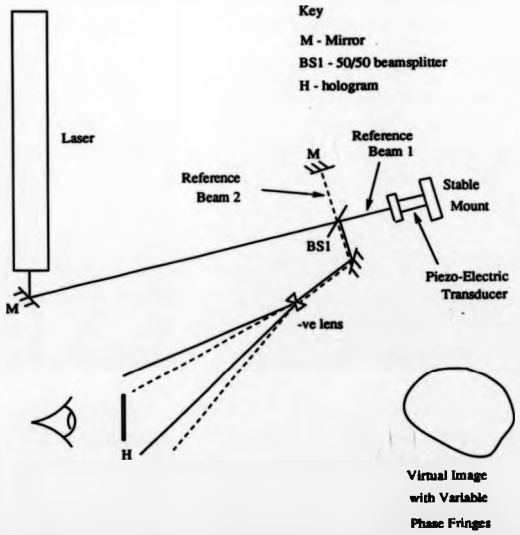


Figure 2.10: Reconstruction System for Dual Reference Beam, Double Exposure Holographic Interferometry

a plane parallel glass plate.

ii) Rotation of a half wave plate in a polarising interferometer.

iii) Translation of a diffraction grating, or equivalantly using an acousto-optic cell.

iv) Stressing a single mode optical fibre which carries the reference beam

The most common technique used is a translatable mirror in the form of a piezo-electric transducer (PZT). These devices are commercially available and can be made to give approximately linear response for displacements up to 1 micron. These devices display a number of problems : non-linear motion with applied voltage, the devices tend to tilt as well as translate, and a good quality interferometer is needed to adequately calibrate the device prior to its use in holographic interferometry. Consequently, care must be taken when large beam cross sections are used because of the possibility of tilt as well as displacement occurring.

A slightly simpler analysis scheme is possible for real time phase stepping analysis, (see figure 2.11). A single reference beam is used with one mirror in the beam path mounted on a PZT, or similar device. An initial exposure is made. After development and repositioning of the hologram, phase stepping can simply be achieved by motion of the PZT. This causes a relative phase change between the two interfering wavefronts : the reference beam which now reconstructs the initial object state, and the sample beam scattered from the new state of the object.

Phase stepping methods have also been applied to time average holography for vibration analysis [17]. The technique uses a PZT for phase stepping at the frequency of excitation of the object. In this way objects with single frequency responses can be quantitatively analysed from time average, or Bessel function, fringes.

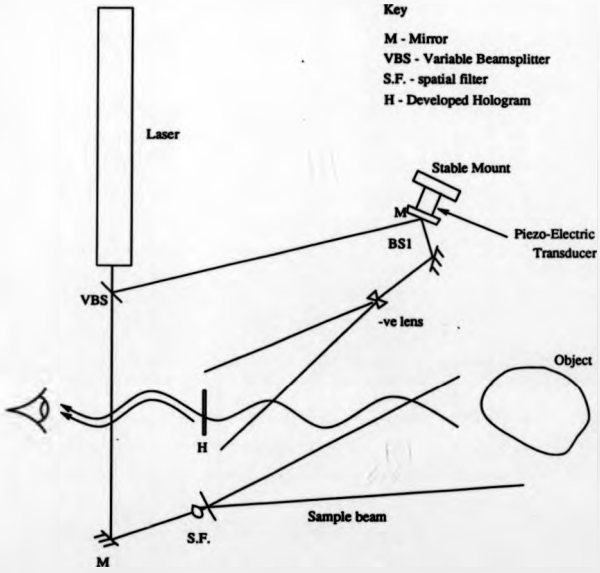


Figure 2.11: Reconstruction System for Phase Stepping, Real Time Holographic Interferometry

A major disadvantage is apparent in the preceding description of holography and holographic interferometry because of the wet processing of the holographic plates which is needed. This inherently limits the time for an interferogram to be reconstructed and the subsequent fringe analysis. The one exception to this is the use of thermoplastic recording materials [15]. However, further research is required to extend the lifetime of these materials, which is currently limited to $\approx$ 100 recordings per thermoplastic plate.

The generally accepted solution is to replace the silver halide recording medium with a television camera and view the interference data (speckle pattern) directly. This type of system has become known as ESPI or TV Holography (TVH). The three dimensionality of holography is purposely sacrificed in order that an image of the object and its speckle pattern are formed onto the imaging sensor. The operation of these systems is described in the following section.

2.4 Electronic Speckle Pattern Interferometry

To observe the interference pattern between the sample and reference beam, the inherent speckle effect must become resolvable. This is achieved in practice by an appropriate aperture and an on-axis reference beam (see equation 2.9). This is shown in figure 2.12 which shows a typical ESPI system [14].

When a diffusely reflecting surface is illuminated, light scattered from the surface is combined with the reference beam at the image sensor, and a speckle pattern is produced. The intensity of an individual speckle, $\mathcal{I}_1$, is given by [14, p.146- 148]

$$\mathcal{I}_1 = I_1 + I_2 + 2\sqrt{I_1 I_2}\cos\psi, \tag{2.57}$$

where I_1 is the intensity of the object beam, I_2 is the intensity of the reference beam, and ψ is the relative phase of the two beams, or speckle phase. If the surface is

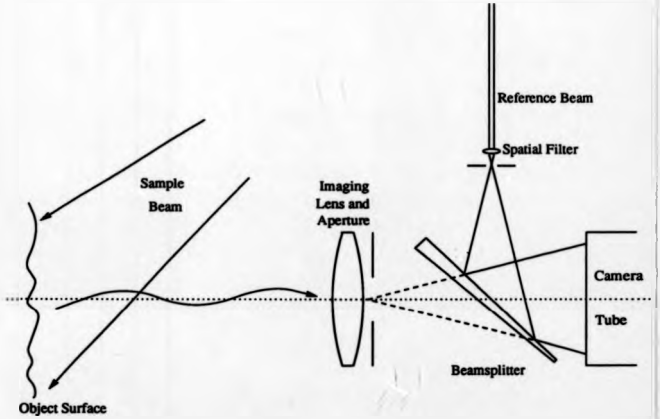


Figure 2.12: Optical System for Electronic Speckle Pattern Interferometry

deformed a second speckle pattern will be formed, $\mathcal{I}_2$ where

$$\mathcal{I}_2 = I_1 + I_2 + 2\sqrt{I_1 I_2}\cos(\psi + \Delta\phi), \tag{2.58}$$

and $\Delta\phi$ is the phase change due to the surface motion, (see section 2.3.1). Each of these speckle patterns is digitised and stored for analysis.

The formation of the fringe patterns in ESPI is fundamentally different from holographic interferometry where the coherent addition of two waves produces the fringes. With ESPI, the fringes are formed incoherently by correlation of the two speckle patterns $\mathcal{I}_1$ and $\mathcal{I}_2$. This may be achieved either by subtraction or addition of $\mathcal{I}_1$ and $\mathcal{I}_2$.

Speckle pattern subtraction is performed by storing the initial speckle pattern $\mathcal{I}_1$. The live camera signal of the deformed object is subtracted from the stored initial state. Regions of the signal with good correlation subtract to give zero, whilst uncorrelated regions give a non-zero value. The resulting video signal, V_s, may be described by [14, p.167-168]

$$\begin{aligned} V_s &\propto \mathcal{I}_1 - \mathcal{I}_2, && (2.59)\\ &= 2\sqrt{I_1 I_2}\left[\cos\psi - \cos(\psi + \Delta\phi)\right], && (2.60)\\ &= 4\sqrt{I_1 I_2}\sin\left(\psi + \frac{1}{2}\Delta\phi\right)\sin\frac{1}{2}\Delta\phi. && (2.61) \end{aligned}$$

This signal is rectified to remove negative values and then displayed on a monitor such that the brightness, B, is proportional to the signal voltage.

$$B \propto \left[I_1 I_2 \sin^2\left(\psi + \frac{1}{2}\Delta\phi\right)\sin^2\frac{1}{2}\Delta\phi\right]^{\frac{1}{2}}. \tag{2.62}$$

If the intensity of the speckles is averaged over a line of constant $\Delta\phi$, then regions of high correlation, $\Delta\phi = 2n\pi$ (n integer) give $B = 0$ (as ψ represents the random speckle phase). Similarly, regions of low correlation $\Delta\phi = (2n+1)\pi$ produce $B \propto \sqrt{I_1 I_2}$. The speckle correlation pattern in figure 1.7 was formed using this process.

The fringe function $\sin^2 \frac{1}{2}\Delta\phi$ is equivalent to $\frac{1}{2}(1 - \cos\Delta\phi)$ which is the same for real time interferometry of a static object.

Speckle pattern addition is performed in a similar manner, producing a video signal, V_a, given by [14, p.168- 170]

$$V_a \propto \mathcal{I}_1 + \mathcal{I}_2, \tag{2.63}$$

$$= 2I_1 + 2I_2 + 4\sqrt{I_1 I_2}\cos\left(\psi + \frac{1}{2}\Delta\phi\right)\cos\frac{1}{2}\Delta\phi. \tag{2.64}$$

To achieve a similar result as in the subtraction case, the signal must be high pass filtered (for D.C. removal) before rectification. The brightness is then given by

$$B \propto \left[I_1 I_2 \cos^2\left(\psi + \frac{1}{2}\Delta\phi\right)\cos^2\frac{1}{2}\Delta\phi\right]^{\frac{1}{2}}. \tag{2.65}$$

This is a fringe pattern of $\cos^2\frac{1}{2}\Delta\phi$ fringes modulated by high frequency speckles. The fringes are of the same period as in the subtraction case but fringe maxima are now minima and vice versa. Addition of speckle patterns is not commonly used due to their higher noise [14]. This can be seen from the initial equation for the sum of the speckle patterns. The occurrence of any bad pixels (bright or dark points in both $\mathcal{I}_1$ and $\mathcal{I}_2$) will add and not be removed by the high pass filter.

Phase stepping methods have been applied to ESPI data in a similar way to holographic fringes. An appropriate phase modulator is inserted into the reference beam path to add discrete phase offsets. This was first shown by Nakadate who used a total of five images [18]. The first image S_0 is of the undeformed object. A set of images $S_1, ..., S_4$ are then captured of the deformed object with reference beam phase offsets of 0, $\frac{\pi}{2}$, π, and $\frac{3\pi}{2}$. Four speckle subtraction fringe images are then formed $(S_i - S_0)^2$, where $i = 1, ..., 4$. The speckle noise term was suppressed by five passes of a 3×3 averaging filter. Four images result with

$$i_n(x,y) = a(x,y).\left(1 - m(x,y).\cos\left(\Delta\phi(x,y) + n\frac{\pi}{2}\right)\right) \tag{2.66}$$

with $n = 0, ..., 3$. These can be solved using the conventional algorithm for four phase stepped images

$$\Delta\phi(x,y) = \tan^{-1}\left\{\frac{i_3 - i_1}{i_0 - i_2}\right\}. \tag{2.67}$$

A different algorithm was implemented by Creath [19] and Robinson [20]. In this work the speckle phase is calculated from the interference of the undeformed object's speckle pattern and the reference beam by using three or four phase stepped images. After deformation of the surface, a second set of phase stepped fields are captured and the phase calculated again. This results in the speckle phase plus the phase change due to the deformation. The two calculated phases are then subtracted to yield the phase change due to deformation. A wrapped phase map is produced containing high speckle noise. Filtering operations are then carried out before phase unwrapping.

Several other enhancements have been made to the concept of ESPI the majority of which aim to reduce the speckle effect in the fringe patterns. These include speckle averaging [21, 22], and speckle phase control [22]. The real time nature of ESPI and the application of phase stepping analysis schemes indicate that these systems will have a large role to play in the quantitative inspection of industrial components.

2.5 Applications of Holographic Interferometry and ESPI

The applications discussed in this thesis include : static deformations, vibrations analysis, aerodynamic compressible flows, and the evaluation of flames. Three problems are evident for some of this work.

i) The quantitative analysis of transient events requires all the information to be captured at two instants in time.

ii) The number of fringes which may occur in a single interferogram of a flame or a turbulent compressible flow can be ≈ 1000 [23]. This cannot be resolved using a 512×512 image sensor and framegrabber.

iii) A pulsed laser is required for the analysis of transient events.

The requirement of a pulsed laser is inherent for all unstable events and can be applied to both ESPI and holographic interferometry [24, 25].

The quantitative analysis of transient events has been demonstrated in holographic interferometry. This follows from the discussion in section 2.3.5 where the phase stepped fringe fields are acquired in the reconstruction stage which is assumed to be time independent. The method does require dual reference beam holographic recording. This can be achieved using high speed optical switching components [26, 27]. In the case where the reference beam sources are close together in space, a simple vibrating mirror may also perform the same function.

An alternative to using high speed optical switching components for phase stepping analysis of transient events has been reported by Kujawinska [28]. In this work a diffraction grating is placed in front of the holographic plate in the recording system. The zero and ± 1 diffraction orders each give rise to holographic images. Between the two exposures, the grating is translated. This produces fixed phase shifts of opposite sign in the ± 1 order fringe patterns. By overlaying each of these images a phase stepping analysis can be performed.

The FFT method can also be applied to the holographic analysis of transient events. All that is required is for one of the interfering wavefronts to be tilted between the exposures. The phase stepping approach is generally preferred over the FFT method, because of the restrictions imposed on the phase gradient which can be analysed (as discussed at the end of the preceding chapter).

The analysis of transient events using ESPI is not obvious from the discussion of the analysis systems given above. This can be seen from the number of images required both before and after deformation of the object under study. Transient event analysis using ESPI has recently been addressed by Kujawinska [29]. The method described is to capture four frames of the undeformed object followed by one frame after applying the deformation, i.e. a reversal of the order proposed by Nakadate. This allows a single event to be examined with reference to some initial condition.

An exact solution for the deformation phase is also achievable from the four phase stepped speckle subtraction images formed in this process. This has two advantages :

i) no averaging is required of the cos fringes; this reduces processing time and removes some errors in the retrieved phase,

ii) the use of an averaging filter limits the rate of change of phase which can be successfully observed in the image.

An alternative approach is the use of a diffraction grating in front of the imaging lens, which produces three spatially separated images of the object on the imaging sensor [29]. Phase stepping is performed in the same way as described above, where the diffraction grating was used to record multichannel holographic data. The number of fringes which can be resolved in the images is thereby reduced when compared with an fringe pattern occupying the whole sensor.

Therefore it is possible for both holographic interferometry and ESPI to provide measurements of transient events. The major limitation for the advancement of ESPI analysis systems is the spatial resolution of the fringe patterns which are obtained. This is especially restricting when the image sensor requires sub-dividing in order to store the phase stepped fringe fields using the grating approach. A significant amount of smoothing is also necessary prior to analysis which further restricts the density of fringes in the original images. The current hardware which is available offers an image resolution of 1024×1024 pixels. This would only be satisfactory for simple phase object applications. In holographic interferometry the density of the fringe patterns in the digitised images can be adjusted during reconstruction. Higher density fringe fields can then be processed by piecing together adjacent regions of digitised data.

The majority of the work in this thesis has been performed using holographic interferometry as this is the only technique which can be used for all the applications listed at the start of this section. The use of ESPI for the investigation of surface deformations (either static or oscillatory) is expected to take over from holographic interferometry because of the real time processing which it offers. Therefore in this

area results are presented from both holographic interferometry and ESPI.

2.6 Thesis Format

The format for the rest of the thesis is described below. The preceding discussion implies that double exposure holographic interferometry with a pulsed laser offers a recording system which can be applied to almost all applications. In general the reconstruction system cannot be identical to the recording system due to industrial constraints. Therefore the phase measurement errors caused by non-identical recording and reconstruction have to be considered to obtain the highest possible resolution. To this end holographic recording and reconstruction arrangements are discussed in the first part of chapter 3. This chapter is centred around three aspects.

i) The presentation of a novel method for obtaining variable phase fringe images from a simple hologram reconstruction arrangement using a single beam. It is then shown how a set of fringe fields at known phase steps can be acquired using this technique.

ii) Automatic methods to take a set of phase stepped fringe patterns, from any recording situation (holographic, ESPI, etc.), and produce a wrapped phase map of sufficient signal to noise for successful phase unwrapping.

iii) An algorithm to automatically produce unwrapped phase maps in a path independent manner despite the presence of noise and aliasing errors in the wrapped phase map.

An analysis of the holographic measurement resolution and accuracy is given in chapter 4. The error sources considered include : single beam hologram reconstruction, the phase stepping technique applied, and environmental effects. Two methods for error compensation are then proposed such that phase measurement accuracies of $\frac{\lambda}{68}$ can be attained for the case of different lasers and optical systems being used for recording and reconstruction. These methods are demonstrated by computer simulations.

Chapter 5 presents the results from several applications of the holographic analysis methods and from ESPI. These are split into two sections : the first considers am-

plitude objects, the second phase objects. The static deformation examples include holographic examples to verify the measurement accuracies attainable, and ESPI cases to illustrate the processing of fringe fields with varying levels of noise. The analysis of phase objects is further split into the study of compressible flows and high temperature burner flames. Quantitative measurements in aerodynamics are an important requirement industrially and one where holographic interferometry has offered some measurements in the past. However this is one area where the development of other techniques, notably particle image velocimetry (PIV), provides a quicker and directly quantitative measurement of relevant flow parameters. This is particularly the case when fully three dimensional flows are studied. The study of burner flames on the other hand is more amenable to holographic analysis and major developments are expected in this field in the future. The work shown here illustrates how this may be brought about through the use of holographic optical elements (HOE's) and the simultaneous formation of two phase stepped fringe fields using a modified interferometric arrangement. The applications of HOE's in industrial holographic interferometry are just beginning to be recognised. Their usefulness relies on efficient chemistry for hologram development and has come about because of the progress made in display holography.

The work is concluded in chapter 6 and the possible avenues for further research are described.

Bibliography

[1] A. A. Michelson, "Studies in Optics", Phoenix books, University of Chicago Press, Third Impression 1962, chapters 1,2,3, and 7.

[2] C. Wykes, in Progress in Optics, ed. E. Wolf, volume 25, p.273-359, 1988

[3] S. G. Lipson, H. Lipson, "Optical Physics", Cambridge University Press, Cambridge, 1969.

[4] C. M. Vest, "Holographic Interferometry", Wiley series in pure and applied optics, 1979.

[5] P. Hariharan, "Optical holography", Cambridge University Press, 1984.

[6] D. Gabor, "A new microscope principle", Nature, volume 161, p.777-778, 1948.

[7] D. Gabor, "Microscopy by wavefront reconstruction", Proceedings of the Royal Society A197, p.454, 1949.

[8] E. N. Leith, J. Upatnieks, "Reconstructed wavefronts and communication theory", Journal of the Optical Society of America, volume 52, p.1123, 1962.

[9] E. N. Leith, J. Upatnieks, "Wavefront reconstruction with continuous tone objects", Journal of the Optical Society of America, volume 53, p.1377-1381, 1963.

[10] E. N. Leith, J. Upatnieks, "Wavefront reconstruction with diffused illumination and three dimensional objects", Journal of the Optical Society of America, volume 53, p.1295, 1964.

[11] Collier, Buckhardt, Lin, "Optical holography", Academic press inc., 1971.

[12] R. L. Powell, K. A. Stetson, "Interferometric vibration analysis by wavefront reconstruction", Journal of the Optical Society of America, volume 55, p.1593-1598, 1965.

[13] R. L. Powell, K. A. Stetson, "Interferometric hologram evaluation and real time vibration analysis of diffuse objects", Journal of the Optical Society of America, volume 55, p.1694-1695, 1965.

[14] R. Jones, C. Wykes, "Holographic and Speckle Interferometry", Cambridge University Press, Cambridge, 1983.

[15] R. J. Parker, "A Quarter Century of Thermoplastic Holography", Society for Experimental Mechanics conference proceedings on Hologram Interferometry and Speckle Metrology, p.217-224, 1990.

[16] K. Creath, "Phase Measurement Interferometry Techniques", in Progress in Optics, ed. E. Wolf, volume 26, p.349-393, Amsterdam, 1988.

[17] K. Oh, R. J. Pryputniewicz, "Application of Electro-Optic Holography in the Study of Cantilever Plate Vibration with Concentrated Masses", Society for Experimental Mechanics conference proceedings on Hologram Interferometry and Speckle Metrology, p.245-253, 1990.

[18] S. Nakadate, H. Saito, "Fringe Scanning Speckle-Pattern Interferometry", Applied Optics, volume 24, number 14, p.2172-2180, 1985.

[19] K. Creath, "Phase Shifting Speckle Interferometry", Applied Optics, volume 24, number 18, p.3053-3058, 1985.

[20] D. W. Robinson, D. C. Williams, "Digital Phase Stepping Speckle Interferometry", Optics Communications, volume 57, p.26-30, 1986.

[21] K. A. Stetson, "Theory and Applications of Electronic Holography", Society for Experimental Mechanics conference proceedings on Hologram Interferometry and Speckle Metrology, p.294-300, 1990.

[22] J. Davies, C. Buckberry, "Applications of a Fibre Optic TV Holography System to the Study of Large Automotive Structures", Society of Photo-optical and Instrumentation Engineers, volume 1162, San Diego 1989.

[23] P. J. Bryanston-Cross, J. W. Gardner, "Holographic Visualisation of a Combustion Flame", Optics and Lasers in Engineering, volume 9, p.85-100, 1988.

[24] R. J. Parker, D. G. Jones, "Industrial Holography the Rolls-Royce Experience", Society of Photo-optical and Instrumentation Engineers, volume 699, p.111-126, 1986.

[25] M. C. Shellabear, F. Mendoza-Santoyo, J. R. Tyrer, "Processing of Addition and Subtraction Fringes From Pulsed ESPI for the Study of Vibrations", Society for Experimental Mechanics conference proceedings on Hologram Interferometry and Speckle Metrology, p.238-244, 1990.

[26] D. P. Towers, T. R. Judge, P. J. Bryanston-Cross, "Vibration Measurements using Dual Reference Beam Holography", Society of Photo-optical and Instrumentation Engineers, volume 1084, p.218-239, 1989.

[27] G. Lai, T. Yatagai, "Dual-Reference Holographic Interferometry with a Double Pulsed Laser", Applied Optics, volume 27, number 18, p.3855-3858, 1988.

[28] M. Kujawinska, D. W. Robinson, "Multichannel Phase-Stepped Holographic Interferometry", Applied Optics, volume 27, number 2, p.312-320, 1988.

[29] M. Kujawinska, A. Spik, "Quantitative Analysis of Transient Events by ESPI", Society of Photo-optical and Instrumentation Engineers, volume 1121, p.416-423, 1989.

Chapter 3

Quasi-Heterodyne Analysis of Fringe Patterns and Automatic Image Processing

3.1 Dual Reference and Dual Reconstruction Beam Holographic Interferometry

The application of phase stepping techniques to double exposure holographic interferometry requires the use of dual reference beams at the recording stage of the hologram. This form of holographic interferometry is applicable to a wide range of applications as described in the previous chapter. Reconstruction is normally carried out with dual beams and phase shifting of the fringe field is achieved with a piezoelectric transducer (PZT) moving a mirror in the path of one of the reconstruction beams. This forms a complex optical system requiring a high degree of stability over the data readout period. A simpler method for producing phase stepped fringe images from a dual reference beam holographic interferogram was sought, using standard engineering components.

A second consideration for a holographic reconstruction system is the sensitivity to phase measurement error sources which are inherent in the arrangement used.

One of the major limitations is due to repositioning errors of the hologram after wet processing (assuming that thermoplastic recording materials are unavailable or not suitable for the application). When this occurs the measured phase of the fringe field no longer represents the deformation of the object alone [1]. The reasons for this are explained below.

In double exposure holographic interferometry the hologram is exposed and then removed for development. Exact repositioning of the hologram in the recording system is then required to form a distortion-free fringe field. This is often not possible when the hologram is produced in an industrial environment, for example a wind tunnel. If the hologram is not repositioned exactly, the effect is equivalent to a change in the reconstruction beam from the original reference beam used to record the object. A change in the beam implies a distortion in the reconstructed fringe field. This can be seen qualitatively from equation 2.1 in the previous chapter which is reproduced here

$$I(x,y) = r(x,y)\bar{r}^*(x,y) + \bar{o}(x,y)\bar{o}^*(x,y) + \bar{r}(x,y)\bar{o}^*(x,y) + \bar{r}^*(x,y)\bar{o}(x,y), \quad (3.1)$$

and represents the two interfering waves (object, $\bar{o}(x,y)$ and reference, $\bar{r}(x,y)$) recorded in a hologram. When the developed hologram is reconstructed with a beam $\bar{r}_p(x,y)$ the virtual image wave $\bar{w}_3(x,y)$ is given by

$$\bar{w}_3(x,y) \propto \bar{r}_p(x,y)\bar{r}^*(x,y)\bar{o}(x,y). \quad (3.2)$$

When $\bar{r}_p(x,y) = \bar{r}(x,y)$ then a perfect reconstruction of the object wave is obtained. In the case where $\bar{r}_p(x,y) \neq \bar{r}(x,y)$, then the virtual image is equal to the object wave multiplied by a term with spatially varying amplitude and phase. In many practical situations the amplitude of the reference wave is constant across the area of the holographic plate. The main difference between the two waves is a spatially varying phase distribution. Therefore $\bar{r}_p(x,y) = r_p e^{-i\psi_p(x,y)}$ and $\bar{r}(x,y) = re^{-i\psi(x,y)}$. Hence

$$\bar{w}_3(x,y) = r.r_p.o(x,y).e^{-i(\psi_p(x,y)-\psi(x,y))} \quad (3.3)$$

and a spatially varying phase change across the object is produced. When this effect is present in two interfering virtual waves it can be seen that the phase of the fringes represents both the object deformation and some function of the difference between $\bar{r}_p(x,y)$ and $\bar{r}(x,y)$. This may also include wavelength differences between the lasers used for recording, e.g. a pulsed ruby laser $\lambda_{Ruby} = 694.3$nm, and reconstruction, e.g. a He-Ne $\lambda_{He-Ne} = 633$nm.

In dual reference beam holographic interferometry inaccurate hologram repositioning, and wavelength differences between recording and reconstruction, have different effects on the wavefronts generated by each reconstruction beam. Hence a distortion of the fringe pattern results. Dändliker et al [1] showed that the error in the phase measurements are reduced by using the two reference beam sources close together in space. The holographic recording systems for separated and close together reference beams are shown in figure 3.1. A Michelson interferometer is used to generate the two reference beam source positions close together in space.

The sensitivity to hologram misalignment is derived below [1]. The analysis considers the formation of a dual reference beam hologram of a single object point O. The direction of the wavefronts in the hologram plane are given by the propagation vectors $\underline{k}$ of magnitude $\frac{2\pi}{\lambda}$. By taking the plane wave approximation to the wave in the hologram plane the phase, $\phi(\underline{x}_h)$, at the point $\underline{x}_h$ is given by

$$\phi(\underline{x}_h) = \underline{k}.\underline{x}_h, \tag{3.4}$$

with respect to the origin of the co-ordinate system, see figure 3.2.

Using this notation, see equation 3.2, the reconstructed virtual image wave is given by

$$\begin{aligned} \tilde{w}_3(\underline{x}_h) &= a(\underline{x}_h).e^{i\left(\omega t - \underline{x}_h.(\underline{k}_p - \underline{k}_r + \underline{k}_o)\right)}, \\ &= c(\underline{x}_h).\bar{o}(\underline{x}_h).e^{i\left(\omega t - \underline{x}_h.(\underline{k}_p - \underline{k}_r)\right)}, \end{aligned} \tag{3.5}$$

where $\underline{k}_o$, $\underline{k}_r$, $\underline{k}_p$ are the propagation vectors of the object, reference, and reconstruc-

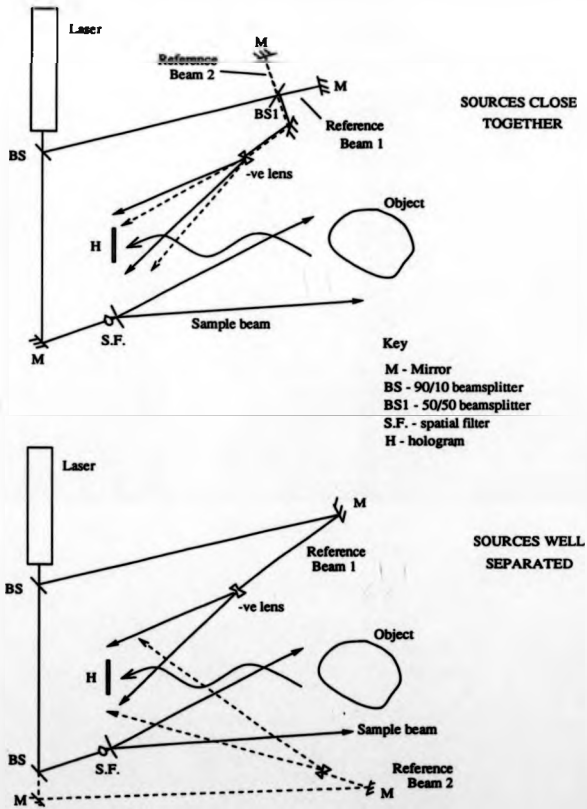
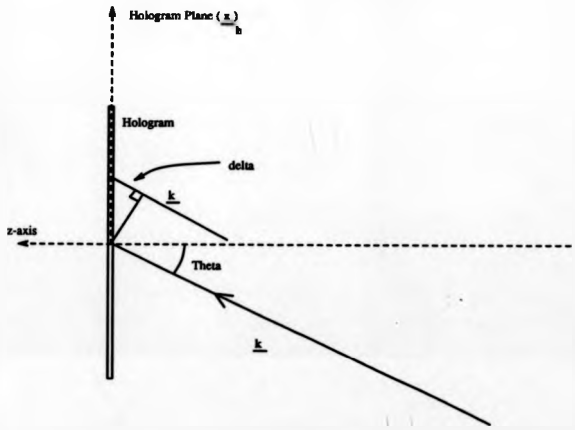


Figure 3.1: Dual Reference Beam Holographic Interferometry with Reference Beam Sources Close Together or Separated



$$\text{delta} = \underline{k} \cdot \underline{x}_h = |\underline{k}| \, |\underline{x}_h| \cos(90+\text{Theta})$$

Figure 3.2: Definitions for Dual Reference Beam Reconstruction Analysis

tion beams respectively, ω is the optical frequency, $\bar{o}(\underline{x}_A)$ is the original object wave in the plane of the hologram, and α, c are constants. This equation models the situation where a wave differing from the reference beam is used for reconstruction. In general, after wet processing, the hologram will not be correctly repositioned, and hence a new vector $\underline{x}'_A$ is introduced to define the new hologram position. This affects the phase in the reconstructed wave, which is now a function of $\underline{x}'_A$. Equation 3.5 is then re- written as

$$\bar{w}_3(\underline{x}_A) = c(\underline{x}_A).\bar{o}(\underline{x}_A).e^{i\left(\omega t - \underline{x}'_A.\underline{k}_p + \underline{x}_A.\underline{k}_r\right)}. \tag{3.6}$$

Considering the case of dual reference beams for recording two object states and two different reconstruction beams, the subscripts 1 and 2 are introduced. The two reconstructed wavefields are then

$$\bar{w}_{3,1}(\underline{x}_A) = c(\underline{x}_A).\bar{o}_1(\underline{x}_A).e^{i\left(\omega t - \underline{x}'_A.\underline{k}_{p1} + \underline{x}_A.\underline{k}_{r1}\right)}, \tag{3.7}$$

$$\bar{w}_{3,2}(\underline{x}_A) = c(\underline{x}_A).\bar{o}_2(\underline{x}_A).e^{i\left(\omega t - \underline{x}'_A.\underline{k}_{p2} + \underline{x}_A.\underline{k}_{r2}\right)}, \tag{3.8}$$

The exponents of these terms represent the extra phase present in the reconstructed waves. The difference between these values is then the extra phase $\psi(\underline{x}_A)$, or error, which is added to the deformation phase required :

$$\begin{aligned} \psi(\underline{x}_A) &= \underline{x}'_A.\underline{k}_{p1} - \underline{x}_A.\underline{k}_{r1} - \underline{x}'_A.\underline{k}_{p2} + \underline{x}_A.\underline{k}_{r2}, \\ &= \underline{x}_A.(\underline{k}_{r2} - \underline{k}_{r1}) + \underline{x}'_A.\left(\underline{k}_{p1} - \underline{k}_{p2}\right). \end{aligned} \tag{3.9}$$

The expression may be simplified by defining :

$$\begin{aligned} \Delta\underline{x}_A &= \underline{x}'_A - \underline{x}_A, \\ \Delta\underline{k}_1 &= \underline{k}_{p1} - \underline{k}_{r1}, \\ \Delta\underline{k}_2 &= \underline{k}_{p2} - \underline{k}_{r2}. \end{aligned} \tag{3.10}$$

Then ψ may be re-expressed as

$$\psi(\underline{x}_A) \approx (\underline{k}_{r1} - \underline{k}_{r2}) . \Delta\underline{x}_A + (\Delta\underline{k}_1 - \Delta\underline{k}_2) . \underline{x}_A, \tag{3.11}$$

by assuming the product of small changes is approximately equal to zero. The vector $\Delta\underline{x}_A$ representing the change in hologram position may be split in to terms for translations $\underline{t}$ and rotations $\underline{w} \wedge \underline{x}_A$ where $\underline{w} = (\Delta\alpha, \Delta\beta, \Delta\gamma)$ defines small rotations about the x, y, and z axes respectively. By examination of equation 3.11 and the definition for $\Delta\underline{x}_A$, it can be seen that the translation of the hologram applies a constant phase change. This will not be seen in the fringe field and hence it can be neglected. Hence equation 3.11 can be written

$$\psi(\underline{x}_A) = [(\underline{k}_{r1} - \underline{k}_{r2}) \wedge \underline{w}) . \underline{x}_A + (\Delta\underline{k}_1 - \Delta\underline{k}_2) . \underline{x}_A. \tag{3.12}$$

The first term of this equation describes the effect of a hologram rotation on the phase of the reconstructed waves. When there is a large angle between $\underline{k}_{r1}$ and $\underline{k}_{r2}$ the sensitivity to any hologram rotation is high. This is the basis for requiring closely separated reference beams in order to minimise errors arising from inaccurate hologram repositioning. The second term of equation 3.12 describes the sensitivity to changes in reconstruction beam from the original reference beam. The equation shows that both these effects are approximately linear with respect to hologram position. From an experimental point of view, the use of closely separated reference beams also has advantages. The two reference beams can be derived simply by using a Michelson system, (see figure 3.1), or just by a slight motion of a mirror in the reference beam path. Thus the optical system is almost as simple as that used in conventional holographic interferometry.

The implication of equation 3.12 is that any change in the reconstruction beam directions from the original reference beams can be compensated for by a rotation of the hologram. This technique has been applied by several authors [2, 3] with some success. However the compensation cannot be guaranteed for all cases. This can be

seen from the vector product in the first term which shows that $(\underline{k}_{r1} - \underline{k}_{r2}) \wedge \underline{w}$ is perpendicular to $\underline{k}_{r1} - \underline{k}_{r2}$. Hence the component of $(\Delta\underline{k}_1 - \Delta\underline{k}_2)$ which is parallel to $\underline{k}_{r1} - \underline{k}_{r2}$ cannot be removed by hologram rotation and ultimately limits the accuracy which can be obtained using these techniques.

The use of closely-separated reference beams causes two further problems when dual reconstruction beams are used (as described above). Firstly the cross reconstructions, given by :

$$\bar{r}_{p2}\bar{r}^*_{r1}\bar{o}_1 \quad , \tag{3.13}$$

$$\bar{r}_{p1}\bar{r}^*_{r2}\bar{o}_2 \quad , \tag{3.14}$$

are superimposed on the primary reconstructions

$$\bar{r}_{p1}\bar{r}^*_{r1}\bar{o}_1 \quad , \tag{3.15}$$

$$\bar{r}_{p2}\bar{r}^*_{r2}\bar{o}_2 \quad . \tag{3.16}$$

It has been shown by Dändliker et al [4] that providing the cross reconstruction waves are laterally separated by a distance larger than the speckle size, no fringes will be formed by these waves. Hence the main problem is a reduction in fringe contrast owing to the presence of the cross reconstructions.

The second difficulty arises in the practical implementation of dual reconstruction beam methods. As the relative phase between the two reconstruction beams directly affects the phase of the interfering light waves, the reconstruction system is susceptible to environmental changes. This is especially relevant when large parts of the beam paths are not common between the two beams. For this reason it is important to minimise the image capture time and maximise stability for dual reconstruction beam systems. Electronic stabilisation methods have been developed by Dändliker for this purpose using the feedback from a photodetector to affect the position of a PZT mounted mirror in one reconstruction beam's path [1].

3.2 Single Reconstruction Beam System for Dual Reference Beam Holography

The use of closely-separated reference beam sources and the problems of dual beam reconstruction systems formed the starting point for the new reconstruction method. When the reference beam sources are closely separated a single reconstruction beam can be used which will form both images of the object and produce a fringe pattern. The fringes may be swept over the object by a change in viewing point through the hologram. This is shown in figure 3.4 as a CCD camera on a linear traverse. Hence a mechanism for applying phase steps into the fringe field is produced. In practice, the motion of the viewing point required to obtain an appropriate phase step, of 90 degrees say, is a few millimetres. Thereby the requirement for micro-translators such as PZT's is removed.

The method is applicable to diffusely reflecting objects. For specularly reflecting objects, the information recorded at a single point on the hologram is derived from a single object point. In this case the single reconstruction beam technique will not work. This is not a serious limitation as holography is normally used specifically because of its ability to interrogate optically rough surfaces.

3.2.1 Analogy to Moire Fringe Patterns

The phase sweeping effect can also be described by an analogy to a phase sweeping method used in moire interferometry [5]. The two gratings in this case are the object deformation fringes which localise on or near the object surface and the reference beam fringes in the hologram plane. Due to the depth between the two gratings a phase sweeping effect is seen when the viewing point is moved. This is analogous to the case when one grating is sheared over another in moire interferometry in order to obtain phase stepped fringe images. An example is shown in figure 3.3. The sequence of photographs was recorded using an automatic wind camera from a moving car. The motion of the car provides the change in viewing point and the railings on either side

of the walk-way form the two gratings.

The analogy correctly illustrates the phase sweeping effect observed, but the fringe formation mechanisms in the two cases are fundamentally different. The moire fringes are formed by the incoherent addition of two similar gratings, whilst the holographic fringes are formed by the coherent sum of the object and reference beam fringe systems.

3.2.2 Mathematical Description of the Phase Stepping Mechanism

The phase stepping mechanism can be understood from a mathematical description of the interfering waves.

Consider a hologram where two object waves $\bar{o}_1$ and $\bar{o}_2$ given by

$$\bar{o}_1(\underline{x}_h) = o_1(\underline{x}_h)e^{-i\Phi_1(\underline{x}_h)}, \tag{3.17}$$

$$\bar{o}_2(\underline{x}_h) = o_2(\underline{x}_h)e^{-i\Phi_2(\underline{x}_h)}, \tag{3.18}$$

are recorded using two reference beams

$$\bar{r}_1(\underline{x}_h) = r_1(\underline{x}_h)e^{-i\Psi_1(\underline{x}_h)}, \tag{3.19}$$

$$\bar{r}_2(\underline{x}_h) = r_2(\underline{x}_h)e^{-i\Psi_2(\underline{x}_h)}, \tag{3.20}$$

respectively, where $(\underline{x}_h)$ defines a point in the hologram plane. When the two waves are reconstructed with a single beam $\bar{r}_p$ the virtual image wave is (from equation 3.2)

$$\bar{w}_3(\underline{x}_h) \propto \bar{r}_p(\underline{x}_h)\left(\bar{r}_1^*(\underline{x}_h)\bar{o}_1(\underline{x}_h) + \bar{r}_2^*(\underline{x}_h)\bar{o}_2(\underline{x}_h)\right). \tag{3.21}$$

In practice, the main effect of a surface movement between the two exposures is to change the phase of the object wave, hence $o_1(\underline{x}_h) \approx o_2(\underline{x}_h) = o(\underline{x}_h)$ [6, chapters 14,15]. A similar assumption for the reference and reconstruction beams may be made $r_1(\underline{x}_h) \approx r_2(\underline{x}_h) = r(\underline{x}_h)$ and to simplify the analysis $\bar{r}_p(\underline{x}_h) \approx \bar{r}_1(\underline{x}_h)$.

Figure 3.3: Phase Sweeping Moire Fringes

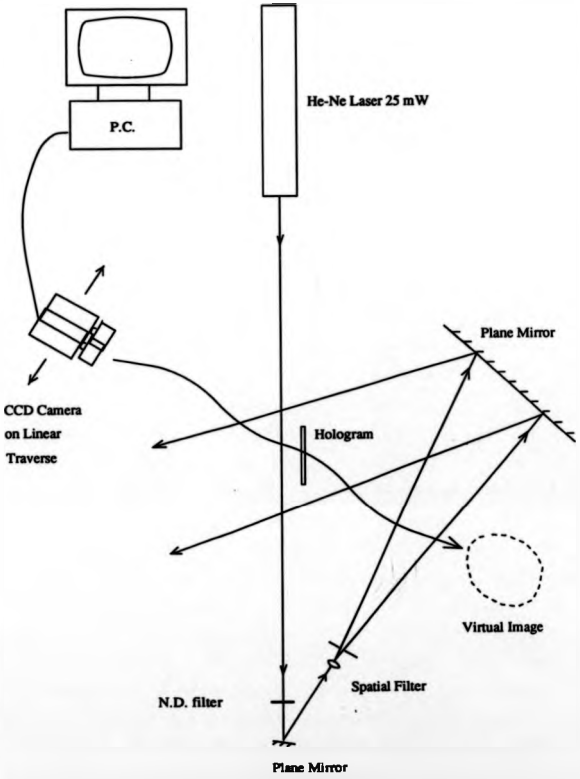


Figure 3.4: Single Beam Reconstruction System for Dual Reference Holograms

Whence equation 3.21 becomes

$$\begin{aligned}
\bar{w}_3 &\propto \bar{r}_1\bar{r}_1^*\bar{o}_1 + \bar{r}_1\bar{r}_2^*\bar{o}_2, \\
&\propto r^2\bar{o}_1 + r^2\bar{o}_2 e^{-i(\psi_{r1}-\psi_{r2})}, \\
\bar{w}_3(\mathbb{Z}_h) &\propto r^2(\mathbb{Z}_h)o(\mathbb{Z}_h)\left(e^{-i\phi_1(\mathbb{Z}_h)} + e^{-i(\phi_2(\mathbb{Z}_h)+\psi_{r1}(\mathbb{Z}_h)-\psi_{r2}(\mathbb{Z}_h))}\right). \qquad (3.22)
\end{aligned}$$

An observer or camera observing this wavefront records the intensity I_{im}. From equation 3.22 :

$$I_{im} \propto (2 + 2\cos(\phi_1 - \phi_2 + \psi_{r2} - \psi_{r1})). \qquad (3.23)$$

Equation 3.23 represents the intensity seen in the reconstructed virtual image. This fringe pattern will be observed providing the reference beam separation at recording is sufficiently small. When the reference beams are further apart the difference in the propagation directions of the two beams will increase. Therefore a lateral shift will occur in the image of the two waves formed by a lens. If the lateral shift approaches the average speckle size in the image then the fringes will start to de-correlate and loose contrast. This effect ultimately limits the separation of reference beam sources which is allowable for the application of the technique. In practice, a separation of $\sim 100\mu m$ at 1m from the hologram will give 20 fringes across a 100mm wide plate. This is reasonably simple to attain and is similar to the requirements given by Dändliker which reduce repositioning errors in the dual reconstruction beam method.

Assuming this condition is satisfied, a fringe pattern will be seen with two components : $\phi_1 - \phi_2$ caused by the deformation of the object, and $\psi_{r2} - \psi_{r1}$ caused by the spatially varying interference phase of the two reference beams in the hologram plane. These terms originate from different physical planes. On reconstruction and viewing at a given point, a particular fringe pattern will be seen according to equation 3.23. If the viewing point is then moved slightly, the second term $\psi_{r2} - \psi_{r1}$ will have changed and hence a different phase of the same fringe pattern will be observed. This assumes that the change in sensitivity vector between the two observation points is negligible. The constraint is easily satisfied by increasing the separation of the reference

beam sources. The spatial variation of the term $\psi_{r2} - \psi_{r1}$ is therefore increased, and a minimal change in sensitivity vector between the observation positions can be produced.

A similar equation can be derived from the analysis by Vest for relating the phase measured $\Phi_1 - \Phi_2$ to the object deformation [7, p.68- 72]. The derivation given in the preceding chapter can be extended to include the effect of dual reference beams, (see section 2.3.1, and figure 3.5).

The phases of the two rays which reach Q from the points P and P1 are given by

$$\begin{aligned} \Phi_1 &= \underline{k}_1 . \underline{r}_1 + \underline{k}_2 . (\underline{R} - \underline{r}_1) + \phi_o + \psi_{r1}, \\ \Phi_2 &= \underline{k}_3 . \underline{r}_3 + \underline{k}_4 . (\underline{R} - \underline{r}_3) + \phi_o + \psi_{r2}, \end{aligned}$$

where ϕ_o is the arbitrary phase of the wave at point O and ψ_{r1} and ψ_{r2} are the phases of the reference beams at Q for the two exposures. In conventional single reference beam holographic interferometry $\psi_{r1} = \psi_{r2}$ and hence

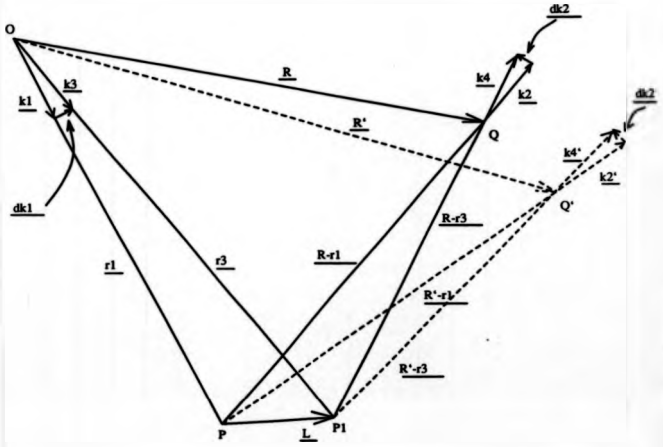
$$\begin{aligned} \delta &= \Phi_1 - \Phi_2, \\ &= (\underline{k}_1 - \underline{k}_2) . \underline{L}. \end{aligned} \tag{3.24}$$

However, in the dual reference beam case the two reference beams are not the same, therefore

$$\delta = (\underline{k}_1 - \underline{k}_2) . \underline{L} + (\psi_{r2} - \psi_{r1}) . \tag{3.25}$$

The fringes in the reconstruction now represent the object deformation modulated by the phase difference caused by the two reference beams. This expression is of the same form as equation 3.23.

The form of the fringe pattern in the hologram plane is dependent on $\bar{r}_1$ and $\bar{r}_2$. When collimated reference waves are used, the phase difference $\psi_{r2} - \psi_{r1}$ will vary linearly across the hologram caused by the formation of Michelson fringes. With



O - Illumination source

P - Initial position of a point on the object

P1 - Deformed position of point P

L - Vector translation of P

Q,Q' - Viewing positions on the hologram

Figure 3.5: Definitions for Dual Reference Beam Sensitivity Vector Analysis

spherical beams the phase difference will also vary across the hologram, e.g. from Q to Q', because of the formation of a Young's fringe pattern. In either case, the fringes can be swept over the object as the reconstructed image is viewed through different parts of the hologram. The separation of the reference fringes formed in the hologram plane determines the camera translation required to produce a 2π phase shift in the fringe field.

Reconstructing a dual reference beam hologram with a single beam will produce a fringe pattern consisting of the object deformation fringes added to the reference beam fringes in the hologram plane. The fringes produced from the two reference beams can be observed by producing a double exposure hologram without deforming the object. When the hologram is reconstructed with a single beam the reference fringes overlay the virtual image of the object. This effect can be seen in figure 3.6. The object is a metal disc which can be deformed in the centre. The reference fringes are slightly blurred owing to the depth between the holographic plate and the virtual image of the disc. The defocussing effect can be minimised by appropriate selection of the imaging lens aperture and position of the camera.

A conventional, single reference beam, interferogram of the loaded disc is shown in figure 3.7. A hologram of the deformed disc was produced using dual reference beams. The single beam reconstruction of this hologram produces the fringe pattern in figure 3.8. The reference fringes appear as a carrier in the image. However, there is no requirement for the fringe pattern to be monotonic as a phase stepping analysis is to be performed using the mechanism described.

The advantages of the single beam system are described in the following points.

i) The technique only requires a single reconstruction beam. Therefore the optical arrangement is simplified to that used for single reference beam interferograms.

ii) The sensitivity to environmental effects on the fringe field is greatly reduced as there is only a single beam in the system when compared with dual beam reconstructions.

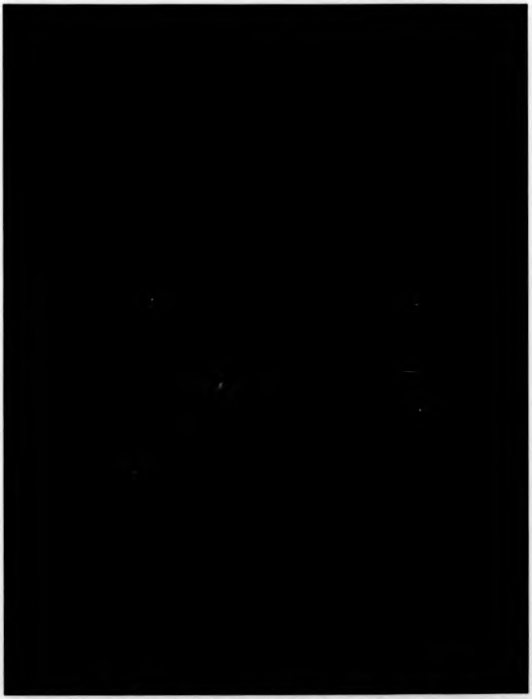


Figure 3.6: Single Beam Reconstruction of Dual Reference Beam Hologram - No Object Deformation

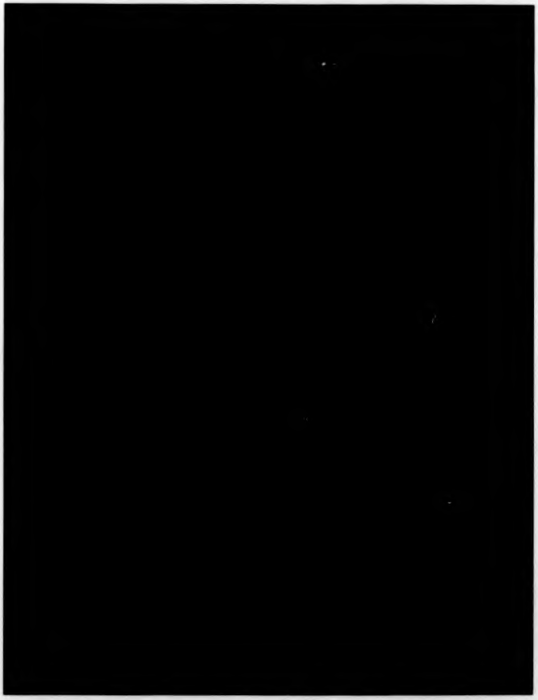


Figure 3.7: Deformation fringes On a Centrally Loaded Disc

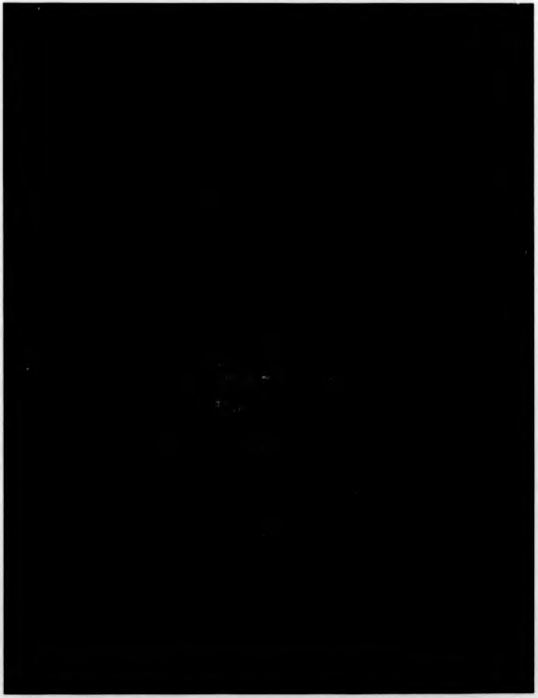


Figure 3.8: Addition of Reference Fringes And Deformation Fringes in Single Beam Reconstruction

iii) No cross-reconstructions are present which reduce the fringe contrast in the dual beam reconstruction system.

iv) Great care is required when dual reconstruction beams are used in order to minimise the effects of repositioning errors and wavelength changes. This results in a time-consuming exercise where the optimisation is often done by eye. This approach was used by Breuckmann et al [2] who quoted a final accuracy of $\frac{\lambda}{50}$.

Although the single beam reconstruction method inherently gives errors in the field to be analysed, i.e. the reference fringes in the hologram plane, the form of the error is known and can be compensated for at a later stage of the data processing. This can be carried out automatically and hence is more efficient as a post-processing operation. Two methods for error compensation are discussed in the following chapter.

3.3 Capture of Phase Stepped Fringe Fields from the Single Reconstruction Beam System

The hologram is reconstructed using the system shown in figure 3.4. A CCD camera is positioned on a linear traverse to look at the virtual image. The video signal from the CCD is received by a digital framestore card (resolution 512 by 512 pixels, 256 grey levels) contained in a P.C. and the output displayed on a monitor. The traverse direction is set approximately parallel to the line joining the two reference beam sources, i.e. normal to the Young's fringes. This produces a maximum change in phase with respect to distance traversed.

The noise produced from the speckle pattern in the reconstructed fringe field incurs a random high frequency intensity fluctuation in the observed fringes. The formation of speckle through an imaging lens is analysed in appendix 1. This results in an expression for the average speckle size in the image plane, d_{sp}, given by [8,

p.56-57] :

$$d_{sp} \approx \frac{2.4\lambda v}{a}, \tag{3.26}$$

where v is the lens to image plane distance, and a is the diameter of the imaging lens aperture. The size of the speckle can be varied by adjusting the lens aperture, large apertures giving a small speckle size and vice versa. Changing the aperture also affects the depth of field of the image. This is important for the single reconstruction beam method described above as fringes are formed in two distinct planes. The localisation of the combined fringe pattern is dependent on the magnitude of the phase gradients present as well as the location of the original phase change. The camera lens is normally focussed on the object and hence the reference fringes appear slightly defocussed. This effect has been studied by Mori et al [9]. Defocussing of the fringes is shown to be analogous to low pass filtering. This is equivalent to the application of an averaging filter over an $n \times n$ window, (see section 1.3.1). The size of the window is determined by the distance of the fringes from the plane of focus. This effect can be utilised to reduce the speckle noise in the fringe field. A limit is reached when the aperture is opened beyond a certain point as the fringes blur out and suffer a marked contrast reduction. From experiments with the single reconstruction beam method, it has been found that apertures smaller than $\sim f^{\oplus}5.6$ must be used. A compromise has to be reached between the speckled appearance of the fringes and the degree of defocussing. Apertures of $f^{\oplus}8$ are typically used in practice and an appropriate neutral density filter inserted into the beam to optimise the available dynamic range of the camera.

As the camera is traversed, the fringes are swept across the object. A translation of the image in the framestore also occurs with a change in traverse position. Most phase stepping algorithms require knowledge of the phase step between the fringe fields. Therefore calibration of the traverse is required to obtain fringe images at a known phase from some reference position. This can be achieved by comparing two fringe images with a 2π radians phase difference. The translation of the image with respect to traverse distance also requires quantifying to allow fringe fields of different

phases to be overlaid within the framestore.

Traverse calibration is achieved by defining a feature in a reference image. The region chosen can be either a part of the fringe field or any well defined object in the field of view. The feature is normally made rectangular for simplicity. This is compared with every possible position the feature can take in a new image using a chi squared test. The camera motion is set parallel to either the columns or rows of the pixels on the CCD. When the motion is horizontal, the image only translates horizontally in the framestore and the position of the feature is confined to a strip in the new image. Each possible position is evaluated and a chi squared value computed using the equation :

$$\chi^2(xpos) = \sum_{j=jst}^{jst+n} \sum_{i=ist}^{ist+m} \frac{(O(xpos+i,j) - E(i,j))^2}{E(i,j)}, \qquad (3.27)$$

where $O(x,y)$ is the grey scale at a point in the new image, $E(x,y)$ is the grey scale of a point in the feature, m,n defines the feature size, and $xpos$ is the position to be tested in the strip of the new image. Chi squared values are obtained for each position the feature can occupy in the new image by incrementing $xpos$ in integer steps. The minimum chi squared value denotes the position of maximum correlation between the feature and its location in the new image. Therefore the pixel shift required to overlay the images in the framestore is defined. The process is then repeated for an image at a new traverse position.

The minimum chi squared value found with respect to traverse distance results in the calibration for the phase change in the fringe field. The minimum chi squared measurement was found to be resolvable to $\frac{1}{100}$th of the traverse distance for a 2π phase shift. This defines a resolution limit in the phase step applied to the fringe field of $\frac{1}{100}$th of a fringe. An example of the chi squared values when the new image is at a 2π phase shift from the reference image is given in figure 3.9. This shows chi squared values plotted against the pixel position in the new image. The image shift correlation is found by the position of the curve minima. The chi squared value at this point is seen to be non-zero implying that perfect correlation is not achieved.

This is due to several factors :

i) the change in sensitivity vector between the two traverse positions,

ii) the change in intensity of the light scattered from the object and sampled at the two traverse positions,

iii) random electronic noise in the camera and digitiser.

The magnitude of the minimum chi squared value is also dependent on the feature size selected.

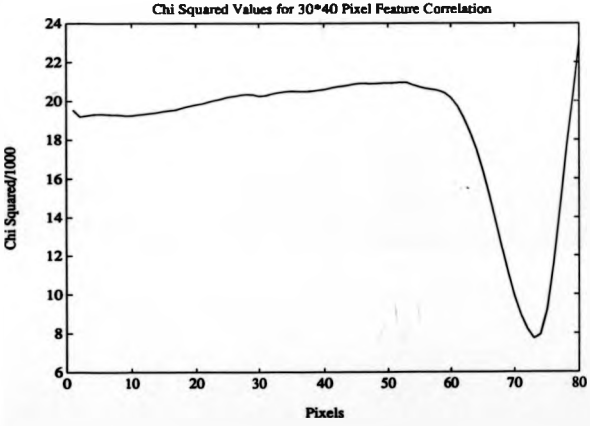


Figure 3.9: Chi Squared Values for Image Shift Correlation

The correlation of the feature in a test image is a computationally intensive process. To increase the execution speed the software was implemented on a transputer

card used in one of the expansion slots of the P.C. The execution time was decreased by a factor of 10 by using the transputer instead of the resident Intel 80286 processor.

The relationship between the phase of the fringe field to the camera translation must be established in order to obtain fringe images at any phase step from a given reference image. Similarly, the pixel shift must be related to the viewing point movement in order to locate any fringe field at the same position in the framestore as the reference image. These relationships can be found by modelling the reconstruction system using geometric optics, see figure 3.10.

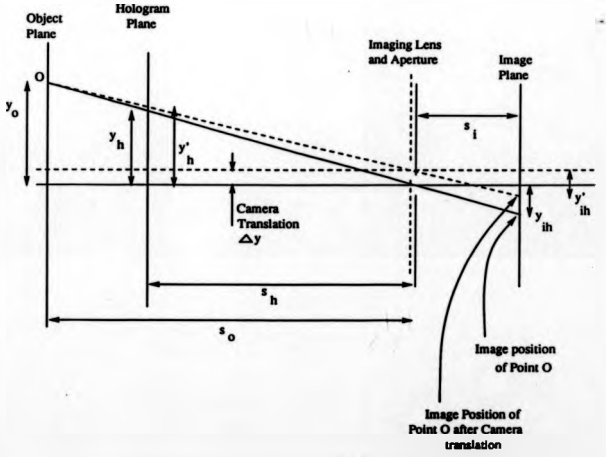


Figure 3.10: Definitions to Calculate Phase Shifting Linearity

The diagram shows two positions of the camera separated by Δy. The camera lens

is modeled by a thin lens and aperture. A single point, O, on the object is considered at a distance s_O from the lens. The hologram plane is defined at a distance s_H from the lens. For the first camera position, the image of O is formed at a distance y_I from the lens axis. The ray paths which originate from O and therefore contribute to the image at y_I describe a circle in the hologram plane. The circle is centred at y_H and the radius is limited by the lens aperture. The average intensity over the circle determines the intensity seen at the image point y_I. The averaging process is synonymous with the blurring of the reference fringes. In this analysis, it is assumed that a sufficiently small aperture is used to give fringes of good contrast. Hence the size of the circle in the hologram plane is small, and the intensity at the point y_H can be used to represent the average intensity over the area of the circle. Therefore the hologram point y_H is imaged to y_{IH} with $y_{IH} = y_I$. For the second camera position, the same object point is considered giving an image at y'_I from the new location of the lens axis. The centre of the circle in the hologram plane is now at y'_H, and the image position is at $y'_{IH} = y'_I$. The difference $y'_H - y_H$ defines the change in the imaged point on the hologram for a camera translation of Δy. This is evaluated below. For the first camera position,

$$\frac{y_O}{s_O} = \frac{y_H}{s_H} \tag{3.28}$$

by similar triangles. A similar expression can be written for the second camera position,

$$\frac{y_O - \Delta y}{s_O} = \frac{y'_H - \Delta y}{s_H}. \tag{3.29}$$

Rearranging equations 3.28 and 3.29 gives

$$y_H \frac{s_O}{s_H} = y_O = \frac{s_O}{s_H}(y'_H - \Delta y) + \Delta y. \tag{3.30}$$

Hence

$$y'_H - y_H = \Delta y \left(\frac{s_O - s_H}{s_O}\right). \tag{3.31}$$

The distance $y'_H - y_H$ defines a change in phase of the reference beam fringes in the hologram plane. When collimated reference beams are used, the phase change is linearly related to the distance between the two imaged points. If the beams are spherical, a Young's fringe pattern results with the distance between neighbouring fringe maxima, δ, is given by [10, p.281-283]

$$\delta \approx \frac{s\lambda}{a}, \tag{3.32}$$

where s is the source to hologram distance, a is the source separation, and λ the wavelength of light. The equation is valid when $s \gg a$. This condition is already required to maintain the accuracy of dual reference beam holographic interferometry. Therefore in both cases the distance $y'_H - y_H$ can be described as a linear function of the reference beam fringe phase. Hence

$$\psi'_H - \psi_H = k\Delta y \left(\frac{s_O - s_H}{s_O}\right), \tag{3.33}$$

where ψ is the reference fringe phase, and k is a constant. This equation is linear with respect to Δy providing the object is planar. The phase change applied is dependent on $s_O - s_H$ with a large object to hologram distance giving a large change in phase for a given camera translation. When a non-planar object is examined, the phase change applied will vary across the hologram according to the change in the term $s_O - s_H$, (from equation 3.33). This effect can be minimised by placing the camera further from the hologram and object.

The linearity of the image translation in the framestore, $y'_{IH} - y_{IH}$, is evaluated below. The magnification of the object point, O, is defined by :

$$M = \frac{y_I}{y_O} = \frac{y_{IH}}{y_O}, \tag{3.34}$$

For the second camera position the magnification is :

$$M = \frac{y_I'}{y_O - \Delta y} = \frac{y_{IH}'}{y_O - \Delta y}. \tag{3.35}$$

However the magnifications in both cases must be equal as the object and image distances are unchanged. Equations 3.34 and 3.35 can be re-expressed as :

$$\frac{y_{IH}}{M} = y_O = \frac{y_{IH}'}{M} + \Delta y. \tag{3.36}$$

Re-arranging gives

$$y_{IH}' - y_{IH} = \frac{s_I \Delta y}{s_O}. \tag{3.37}$$

This expression relates the image translation in the framestore to the camera movement. For planar objects the factor $\frac{s_I}{s_O}$ implies that the pixel shift is constant across the image.

The majority of the examples given in this thesis are of planar objects (the case of non-planar objects is considered in the following chapter). Therefore, from equations 3.33 and 3.37, the pixel translation and phase step are linear functions of the camera movement. The camera translation found from the calibration position (giving a 2π phase step) can therefore be used to obtain an image at any required phase step by linear interpolation. This fringe field can be overlaid on the reference fringe pattern by linear interpolation of the pixel shift for a 2π phase step.

This procedure was first described by the author in the paper [11]. The accuracy of the method is dependent on the fringe period in the traversed direction. In the vibration data presented in figure 1.6 the minimum fringe period is 20 pixels. Consequently, when the images are shifted to the nearest pixel the accuracy of the phase measurements is limited to $\frac{1}{20}$th of a fringe.

This constraint is removed by incorporating linear interpolation of the feature grey scale values down to $\frac{1}{10}$th of a pixel [12]. A correlation performed at sub-pixel resolution is justified because a relatively large number of pixels are being used to form the correlation sample, typically $> 30 \times 30$ pixels. The resolution limit is then

caused by electronic noise in the framegrabber and CCD which gives a frame-to-frame variation in pixel grey levels. The sub-pixel shift fraction which can be resolved has been determined experimentally by correlating two images grabbed at the same traverse position. The pixel shift between the images is found from the position of the minimum chi squared value and defines the repeatability of the correlation. In practice, repeatable results are achieved down to $\frac{1}{4}$ of a pixel. The corresponding resolution for the case of a fringe period of 20 pixels is $\frac{1}{80}$th of a fringe. A graph showing the difference in chi squared values obtained by image correlations to 1 pixel and to $\frac{1}{4}$ pixel are shown in figure 3.11. The curve for $\frac{1}{4}$ pixel resolution shows a reduced minimum value implying a more accurate correlation.

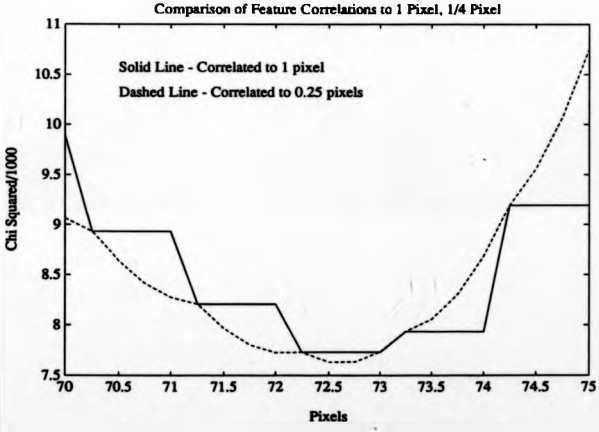


Figure 3.11: Comparison of Chi Squared distributions for Integer Pixel and 1/4 Pixel Resolution Image Translation

Further improvements in resolution are attainable by averaging several frames

of data from a single traverse position. This process will tend to average out the effects of random electronic (or shot) noise. The procedure requires the accumulation and storage of many images. This has not been practical in the past because of the limitations in computer processing power available. The introduction of image processing cards with video rate arithmetic functions however means that 25 frames can be averaged in 1 second without requiring intermediary disc storage [13]. This should reduce the effects of shot noise by a factor of $\frac{1}{\sqrt{N}}$ where N is the number of frames averaged [14, p.347].

The pixel shifting process has the further effect of decreasing the effective area of the sensor. This is not found to be significant as the requirement for negligible change in sensitivity vector implies that the total camera traverse distance is kept to a minimum. Therefore the translation in the framestore is also minimised. A typical experiment performed would result in a total image translation of ~ 45 pixels. Hence the active area of a 512×512 pixel framestore is reduced by only 9%.

3.4 Pre-Processing Of Fringe Data

This section describes the approach used to apply filters to the fringe fields automatically. This forms the first part of the fringe analysis image processing routines. The fundamental requirement of a pre-processing algorithm is to produce fringe data of suitable quality so that subsequent routines are able to extract the phase information automatically. This applies to all fringe analysis methods : fringe tracking, phase stepping, and the Fourier transform method. A second requirement is for the routines to execute automatically, in a minimal time, and without over filtering the data. Over filtering can cause distortions of the phase information and a loss of fringe contrast especially for dense fringes.

In the ideal case no image preparation would be required. However, the examination of diffusely reflecting objects inherently produces speckle noise in the fringe field. In ESPI fringe data, the speckle size is normally larger than that for holography and hence the noise content of the data is higher. Therefore most holographic data and all ESPI images require some filtering before analysis. A compromise must be obtained between the amount of local data lost through filtering whilst maintaining a good solution from the subsequent analysis.

A large amount of work has been done on methods for filtering fringe data as described in the first chapter, (see section 1.3.1). In most cases, the techniques have been developed specifically for a given type of fringe field : interferometric, holographic, or ESPI. As both holographic and ESPI fringe fields are presented in this work, a common filtering algorithm is required to manage both types of data. The algorithm must therefore be able to execute automatically over a range of different types of fringe data.

A spatial domain filter has been chosen as opposed to a frequency domain bandpass approach because of the increased execution speed. The actual filter used in this work is a 3×3 pixel averaging operator. This allows a relatively small amount of filtering to be applied and the number of iterations provides an effective control on the final image quality.

The amount of filtering applied is governed by the size of the kernel used, in this case 3×3 pixels, and the number of iterations of that kernel. By incrementing the number of passes of a filter the quality, or signal-to-noise ratio, can be modified. This process can be controlled by evaluating a measure of the signal-to-noise ratio after each iteration of the filter. This measure is then used to determine if more iterations of the filter are needed or the optimum fringe field quality has been reached. This allows the filtering process to be adapted to the type of data being analysed, automatically applying the optimum degree of filtering required. This method was described by Towers et al [12].

In section 1.3.1 it was shown that a qualitative examination of the FFT of a pixel raster can be used to estimate the signal-to-noise ratio of the fringe data [15]. The high frequency components represent noise, due to laser speckle and the electronics in the CCD and framegrabber, whilst the lower frequencies represent the desired information about the fringe separations. The measure of the signal to noise ratio can be found from the amplitude spectrum computed over a section of the image. The validity of the signal-to-noise parameter may be evaluated against the results from the phase stepping analysis, and the unwrapped phase data.

The method used to estimate the signal-to-noise ratio from the amplitude spectrum has a significant effect on the form of the results produced. Ideally the signal-to-noise parameter should show a definite point at which the optimum fringe quality has been attained. It is essential for automatic pre-processing that this point is detectable by software. A value in the amplitude spectrum is chosen to define the level above which the amplitude is considered to be signal, and below is considered to be noise. Two approaches have been considered : the mean of the amplitude spectrum, and the median. A plot showing the signal-to-noise ratio calculated by these two methods against average iteration is shown in figure 3.12 for a holographic image. The results are calculated with a cos^4 window across the input data. The average of the spectra for 10 adjacent columns has been taken. This process reduces the effect of any random variations from column to column. The signal-to-noise calculated with respect to the average does not produce a clear indication of an optimum point. Some in-

crease is seen but there are other variations present which make the curve difficult to interpret. The median result shows a large increase in signal-to-noise which flattens out after 3 iterations. This curve shows a large initial increase in the signal-to-noise ratio and reaches a maximum value at the fourth iteration.

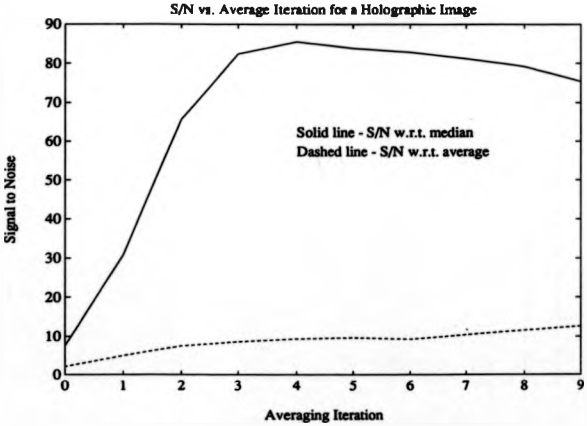


Figure 3.12: Comparison of Methods to Calculate Fringe Field Signal to Noise Ratio

This analysis relies on the quantisation of the image data before the signal to noise is evaluated. Otherwise the high frequency terms in the spectra become approximately zero and the signal to noise rises exponentially. The use of the median in this case becomes equivalent to a bandpass approach where the high frequency half of the spectrum is classified as noise, and the lower half the signal. This analogy can be used to explain the decay in signal to noise with respect to the median which is seen after the fourth iteration in figure 3.12. The first iterations of the filter affects the high

frequency terms in the spectra. To prevent this part of the spectra reducing to zero, the data is quantised back into 8 bits prior to calculating the signal-to-noise ratio. Further iterations then produce high frequency terms with approximately constant amplitude. Further averaging, however, also tends to reduce the amplitude of the signal part of the spectra implying that the data is over filtered. This phenomena is shown in figure 3.12 as a decay in the signal-to-noise ratio parameter.

The variation in the signal-to-noise ratio parameter versus average iteration has been verified against the results from the phase stepping analysis and the phase unwrapping algorithms. In both cases, a variable must be found that is influenced by the quality of the fringe data. One of the stages in phase stepping analysis involves the pixel by pixel calculation of the phase step between the fringe fields. The results are affected by noise in the fringe field and hence the variation in the phase step can be used to indicate the effect of filtering. The phase unwrapping code gives a result for the fraction of the fringe field which has been successfully analysed. However, this result is not a simple function of the fringe quality. Large changes in the area solved can be caused by a small region of the field containing high noise and vice versa. Hence the results from the phase stepping analysis are used for a comparison.

The distribution of the phase step angle calculated over a large fraction of the image is indicative of the fringe field quality. The percentage of points producing a phase step within a fixed range of the mean is calculated for a region of the fringe field. The range defines the points which are considered as well conditioned. In practice, the same form of the results is produced for different values of the phase step range. In the results shown here, the range is set to 20 degrees from the mean. A plot of the percentage of points satisfying this criterion against average iteration is shown in the upper half of figure 3.13 for the same holographic data examined in figure 3.12. The signal-to-noise ratio is repeated in the lower plot to allow a direct comparison to be made. The form of both plots is similar. The optimum fringe quality occurs at the second pass of the averaging filter from the phase step calculation data. Therefore, the signal-to-noise ratio parameter should give a peak at the second pass of the averaging filter. Hence the dashed line in the lower plot has been calculated

which represents the slope of the signal to noise. This curve shows a peak at the second iteration which corresponds with the maxima in the upper plot.

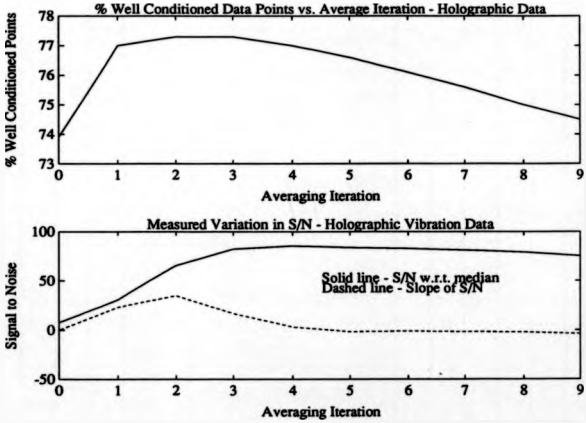


Figure 3.13: Comparison of Measured and Expected Signal to Noise Variations - Holographic Data

In practice, for automatic image pre-processing, the maximum slope point of the signal-to-noise ratio is easily detectable in software. This indicates an optimum number of iterations for the filter. However, there is negligible change in the image area solved between the result for the indicated optimum iteration and the previous pass. Therefore, to reduce the execution time for the filtering process, the actual number of iterations used is one less than the iteration where the slope of the signal-to-noise ratio reaches a maximum.

A similar plot is shown in figure 3.14 for an ESPI image. This would indicate an optimum point at the fourth averaging iteration, the slope of the signal-to-noise

ratio reaching a maxima at the fifth iteration. The upper plot, of the percentage of points producing a phase step within 20 degrees of the mean, does not indicate any optimum point, only a rising level which slowly flattens out. In this case, it becomes an arbitrary decision as to the optimum point. This is a consequence of the higher noise level found in raw ESPI images when compared with the holographic data. The results from the phase unwrapping code indicate that between three and five iterations of the filter produce the maximum area which is correctly solved in the fringe field. Further iterations of the averaging filter gives no increase in the image area solved.

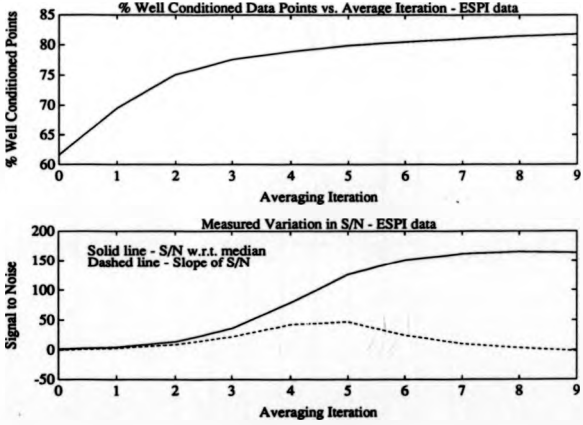


Figure 3.14: Comparison of Measured and Expected Signal to Noise Variations - ESPI Data

This procedure defines an automatic approach to assess the quality of the fringe data such that the maximum amount of phase data can be extracted. The signal-to-noise ratio calculated with respect to the median of the amplitude spectrum provides

a simple means of controlling the amount of smoothing to be applied. This measure can also be used with different types of filter, such as the spin filter described in section 1.3.1. Hence the approach also provides a method with which to evaluate the performance of filters with respect to a common measure. The technique is simple to apply in software and executes in a short time compared with the filter itself.

The pre-processing algorithm has been described specifically with reference to the phase stepping analysis technique. However, similar results can be expected from other analysis procedures. This is verified from the fringe tracking results shown in the introduction, section 1.3.2. The tracking of holographic fringe data was presented showing errors when the raw fringe field was processed, and correct tracking after the application of a single pass of a 3×3 averaging filter. The same requirement for a single pass of the filter has been shown to give the maximum number of well conditioned data points in phase stepping analysis. This indicates a common requirement for the signal-to-noise ratio of a fringe field data such that any analysis technique can be applied successfully.

The one exception where this procedure would not be used in practice is for the FFT analysis method. Here the formation of the wrapped phase map requires the FFT for each raster in the image to be calculated. Hence there is no time penalty for using frequency domain filtering. The limits of the bandpass filter define the allowable frequencies in the final fringe field. The inverse FFT then gives the desired fringe pattern.

3.5 Application of Phase Stepping Algorithms

A set of phase-stepped fringe fields can be captured using the image reconstruction system described in section 3.3. Each image is translated in the framestore to a resolution of $\frac{1}{4}$ pixel such that the fringe fields occupy the same space in the framestore. The fields are then filtered as described in section 3.4 and normalised prior to the application of the phase stepping algorithms.

There are many phase stepping algorithms which have been presented for the

combination of a set of phase-stepped cosinusoidal fringe images. The fringe fields are used to calculate a wrapped phase map by means of an arc tan function [16, 17, 18]. The wrapped phase map contains values for the phase of the fringe field, but wrapped into the interval $-\pi$ to π. As described in the introduction all these algorithms are based on a set of simultaneous equations which may be summarised to :

$$i_n(x,y) = i_1(x,y) + i_2(x,y)\cos\left(\phi(x,y) + \theta_n\right), \tag{3.38}$$

where $i_n(x,y)$ is obtained from the grey scale at a pixel in the digitised image, $i_1(x,y)$, $i_2(x,y)$ are constants, $\phi(x,y)$ is the required phase measurement, n is an integer denoting the image number, and θ_n is the phase step of the nth image. This equation can be seen to have the same form as equation 3.23 with $\phi = \phi_1 - \phi_2$, and $\theta_n = \psi_{r2} - \psi_{r1}$. Experimentally the simplest approach to capture the set of fringe fields is to use a constant phase step between consecutive fields, i.e. $\theta_n = n.\theta$. This implies translating the camera by equal increments, a process which could readily be automated using a computer controlled traverse. This assumes that the phase stepping mechanism is linear, which has been verified in section 3.3. Images captured at equal traverse distances apart must therefore have an equal phase step between them.

After translating the images to occupy the same position in the framestore, the data at each pixel is derived from different speckles in each of the images. This is not the case in conventional systems when the relative position of the CCD camera and reconstructed wavefront does not change. Therefore, the modulation of corresponding image points in the different fringe fields may vary. This effect is reduced by ensuring that the imaging lens aperture is made as large as possible. The speckle size is then reduced and more speckles will be imaged onto each element of the photosensitive detector. This improves the quality of the fringe data as more speckles are being averaged and the intensity pattern becomes closer to a pure cosinusoidal fringe function, see equation 2.26, section 2.3.1.

Two phase-stepping algorithms have been considered. Firstly, using three images

with equal phase steps between the fringe fields where the phase step is known from the traverse calibration data. Secondly, the use of four images with an unknown but constant phase step between the images. Each approach is discussed below.

3.5.1 Three Image Phase Stepping Analysis

The unknowns in the equation 3.38 are i_1, i_2, ϕ, and θ, with ϕ being the required parameter. The traverse calibration data can be used to calculate the phase step θ by linear interpolation. Therefore three equations are sufficient to eliminate i_1 and i_2 and determine the phase ϕ representing the object deformation. The algebraic manipulation of the equations is presented in appendix 2 to give :

$$\phi = \tan^{-1}\left(\frac{(i_{2\theta} - i_0)(\cos\theta - 1) - (i_\theta - i_0)(\cos 2\theta - 1)}{\sin\theta\,(i_{2\theta} - i_0) - \sin 2\theta\,(i_\theta - i_0)}\right). \tag{3.39}$$

This equation is calculated for each point (x, y) in the field.

3.5.2 Four Image Phase Stepping Analysis

Adding a fourth image to the system of simultaneous equations implies that the phase step θ may also be calculated. This procedure eliminates the need for traverse calibration. The algebraic manipulation is given in appendix 2, producing an equation for the phase step :

$$\theta = \cos^{-1}\left(\frac{i_0 - i_\theta + i_{2\theta} - i_{3\theta}}{2(i_\theta - i_{2\theta})}\right). \tag{3.40}$$

Again the equation can be computed for each pixel in the image. These results are averaged over the fringe field to eliminate the effects of noise. The mean phase step and three consecutive images are then used in equation 3.39 to determine the deformation phase.

The use of four images has significant advantages over three as an automatic image processing system results. In practice, manual traverse calibration can take several hours, and a significant amount of computer time for chi squared calculations. The second method only requires the capture of four equi-spaced images from the

traverse. The computational overhead for the phase step calculation is approximately 30 seconds for 40000 image points, evaluated on a SUN-4 processor. This is not considered to be significant.

The use of four images at any phase step implies that a feature of the fringe field cannot be used to calculate the pixel shift between the images. The pixel shift can only be determined using a fringe feature if the two images are known to have a phase step of 360 degrees between them. Therefore in the general case of four images at any phase step, a well-defined object with no fringes is needed to calculate the pixel shift between the first and fourth images. The other fields are then translated by linear interpolation of the shift for the fourth image.

Some care needs to be exercised in all phase stepping analysis schemes in order to maintain the potential accuracy of the method. If the phase step is small, the difference in grey scale between corresponding points in the phase stepped fringe fields will also be small. Equations 3.39 and 3.40 both rely on the accuracy of intensity (or grey scale) differences between the images. Hence the results will contain large variations if the magnitude of the speckle noise rises above the grey scale difference of corresponding points in the fringe fields. This effect is evident with both analysis algorithms. The accuracy of the analysis is therefore dependent on the phase step angle and noise content of the image.

The wrapped phase fringes, see figure 1.16, are represented by a graduation in intensity from black at one edge to white at the other. One fringe width indicates a fixed displacement related to the wavelength of light used and the sensitivity vector, see section 2.3.1.

The procedure described here uses the same equations as presented by Kreis [19]. However in this paper, the phase step calculation was performed as a verification only. Here the calculation of the phase step is used to eliminate the requirement for the phase stepping mechanism to be calibrated and thereby automate this stage of the analysis by image processing.

3.6 Phase Unwrapping

The phase map which results from the application of the phase stepping algorithms contains the required phase data at every point in the image. However, this data is not in a form suitable for analysis by non-specialists due to the incorrect representation of phase values greater than 2π which become wrapped into the range 0 to 2π. The data must therefore be unwrapped before presentation to the user or for subsequent calculation. Whilst the trained human brain can make sense of this kind of phase map, an automatic software analysis must use artificial intelligence methods to 'know' good data from bad and then make sense of the map as a whole. The methods for phase unwrapping were reviewed in chapter 1. From this discussion three conclusions can be drawn.

i) The phase data must be unwrapped in a path-independent manner (otherwise different answers will be found for the phase between two points dependent on the route taken through the image).

ii) Single pixel noise can be detected by either a 2×2 pixel consistency test, or a calculation for the modulation depth of the fringes.

iii) The presence of fringes with high spatial frequencies can lead to violation of the sampling theorem. This introduces an aliasing-type discontinuity in phase values which may be spread over a large region of the image. The fringe which gives rise to the discontinuity may then become broader in a different part of the image, see figure 3.15. In this case, the phase unwrapping process cannot remove all phase discontinuities. If the discontinuity is removed in the region where it exists, the phase data will not be consistent where the fringe becomes broader and is sampled correctly.

Providing all occurrences of inconsistent data at the pixel level are detected correctly, any path-independent phase unwrapping algorithm (cellular automata, minimum weight spanning tree (MST) etc.) will produce correct results. However, when aliasing errors are present the third point implies that no path-independent solution

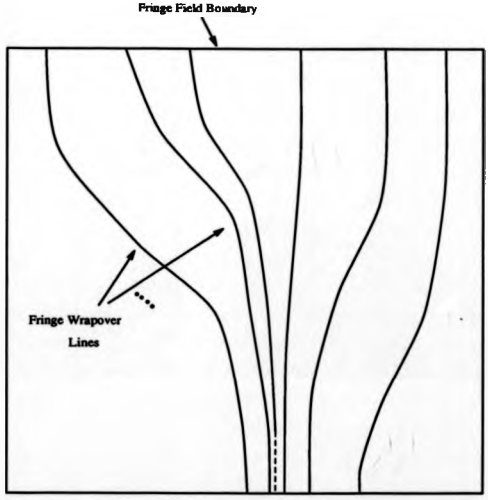


Figure 3.15: Schematic Diagram of a Wrapped Phase Map Showing Fringe Wrapover Lines

exists. Therefore the best result which can be achieved is to isolate the sources of discontinuity and reduce their influence on the rest of the field. In this way the algorithm will fulfill the fundamental requirement to produce meaningful data over the majority of the field despite the presence of bad data and discontinuities.

The approach which has been developed introduces two scales to the fringe analysis. The first is at the pixel level and is equivalent to the cellular automata method of Ghiglia [20, 21]. A second scale is introduced by considering square regions of pixels which have been called tiles. This was initiated and developed by Judge [15, 12, 22].

The tiles are solved by applying the MST technique on a pixel-by-pixel basis. The edge weights are determined by unwrapping the phase along rows and columns. Wrapover points are found using an edge detector which locates phase changes which are greater than π between two pixels. The threshold of π is derived from the Nyquist sampling theorem, (see section 1.3.6). The weights are defined to be the difference in the unwrapped phase values from pixel to pixel and are calculated for the four nearest neighbours to a pixel. The edge weights are then used to construct the MST which defines a path by which the points are joined together. This path is followed with the absolute phase at each point being determined by the unwrapped phase difference between neighbouring points. The result after this stage of processing is a set of tiles containing unwrapped phase data relative to a point within the tile.

An intelligent approach to joining the tiles together is then needed. The set of tiles are considered as large scale pixels and an MST is constructed to join them together, thereby producing an unwrapped phase map over the whole image. The selection of the tile-to-tile weighting factors is important in order to minimise the effects of aliasing type discontinuities. The weight is made up from three factors :

i) the spatial frequency of the fringes (averaged between the two tiles),

ii) the fit between the edges of neighbouring tiles,

iii) the number of low modulation points in the tiles.

All the factors are calculated such that a low weight implies good data or widely spaced fringes. The last two factors relate the local quality of the data. The weighting factor formed from the spatial frequency of the fringes ensures that tiles in a region where there is a slow change in phase are connected first. This is important as aliasing discontinuities tend to be present in regions of closely-spaced fringes.

By assigning tiles with high phase gradients to possess large weights, the regions with aliased fringes are kept to the ends of the MST. Therefore aliased data cannot affect the rest of the tree and the majority of the fringe field is therefore unwrapped correctly. Similarly, any bad data in a tile will affect the phase at a tile's boundary and the weight for that tile will be high. The tile will therefore be linked to its neighbours at the end of the tree, isolating the region of bad data.

These processes for restricting the influence of bad data in a fringe field will be discussed with respect to a number of applications in chapter 5. One of the examples shows a fringe field which is spread over five objects. The boundaries between these objects define paths which should not be used for phase unwrapping. In practice, the presence of multiple objects is dealt with in one of two ways. If the objects are sufficiently separated then the pixels between them fail due to their low modulation. When this does not occur the object boundary can contain sections where the fringes are well defined or blurred out. This can be seen to be similar to an aliasing discontinuity which affects some parts of the field and not others. Therefore it is handled by the algorithm in the same way, forcing the discontinuity in the unwrapped phase to the ends of the tree which are the tiles at the blurred section of the object boundary.

In both pixel and tile level unwrapping, the MST is formed by a modified version of Prim's algorithm [23, p.61- 64]. The operation of the code is best appreciated from a discussion of the results which are obtained. The applications are presented in chapter 5, including both holographic and ESPI examples, and the formation of the solution is described in each case. The finer details of the codes operation are described in reference [12, 22]. The execution time of the phase unwrapping code is typically 2 minutes on a SUN-4 processor for an 512 × 512 pixel image.

Bibliography

[1] R. Dändliker, "Heterodyne Holographic Interferometry", in Progress in Optics, volume 17, ed. E. Wolf, p.1-84, North Holland, Amsterdam, 1980.

[2] B. Breuckmann, W. Thieme, "Computer Aided Analysis of Holographic Interferograms using the Phase Shift Method", Applied Optics, volume 24, number 14, p.2145-2149, 1985.

[3] T. A. W. M. Lanen, "Digital Holographic Interferometry in Flow Research", Optics Communications, volume 79, number 5, p.386-396, 1990.

[4] R. Dändliker, R. Thalmann, J. F. Willemin, "Fringe Interpolation by Two-Reference-Beam Holographic Interferometry : Reducing Sensitivity to Hologram Misalignment", Optics Communications, volume 42, number 5, p.301-306, 1982.

[5] A. Boehnlein, K. Harding, "Field Shift Moire, A New Technique for Absolute Range Measurement", Society of Photo-optical and Instrumentation Engineers, volume 1163, p.2-12, 1989.

[6] P. Hariharan, "Optical Holography", Cambridge University Press, Cambridge, 1984.

[7] C. M. Vest, "Holographic Interferometry", Wiley Interscience series, New York, 1979.

[8] R. Jones, C. Wykes, "Holographic and Speckle Interferometry", Cambridge University Press, Cambridge, 1983.

[9] K. Mori, Y. Nakano, K. Murata, "Fringe Pattern Analysis using Hybrid Image Processing", Applied Optics, volume 29, number 11, p.1646-1651, 1990.

[10] E. Hecht, A. Zajac, "Optics", Addison-Wesley publishing company, 1974.

[11] D. P. Towers, T. R. Judge, P. J. Bryanston-Cross, "The Analysis of Holographic Fringe Data using the Dual Reference Approach", accepted for publication in Optical Engineering, issue April 1991.

[12] D. P. Towers, T. R. Judge, P. J. Bryanston- Cross, "Automatic Interferogram Analysis Techniques Applied to Quasi-Heterodyne Holography and ESPI", Optics and Lasers in Engineering, special issue 'Fringe Pattern Analysis', accepted for publication, 1991.

[13] Matrox product specification for the MVP-AT framegrabber and image processor card, Matrox (UK) Ltd., Cherry Orchard West, Kembrey Park, Swindon, SN2 6UP.

[14] C. Wykes, in Progress in Optics, ed. E. Wolf, volume 25, p.273-359, 1988

[15] D. P. Towers, T. R. Judge, P. J. Bryanston-Cross, "A Quasi-Heterodyne Holographic Technique and Automatic Algorithms for Phase Unwrapping", Society of Photo-optical and Instrumentation Engineers, volume 1163, 1989.

[16] P. Hariharan, "Quasi-heterodyne Hologram Interferometry", Optical Engineering, volume 24, number 4, p.632-638, 1985.

[17] R. Dändliker, R. Thalmann, "Heterodyne and Quasi-heterodyne Holographic Interferometry", Optical Engineering, volume 24, number 5, p.824-831, 1985.

[18] W. Jüptner, T. H. Kreis, H. Kreitlow, "Automatic evaluation of holographic interferograms by reference beam phase shifting", Society of Photo-optical and Instrumentation Engineers, volume 398, p.22-29, 1983.

[19] T. M. Kreis, "Quantitative Evaluation of Interference Patterns", Society of Photo-optical and Instrumentation Engineers, volume 863, p.68-77, 1987.

[20] D. C. Ghiglia, G. A. Mastin, L. A. Romero, "Cellular-Automata method for Phase Unwrapping", Journal of the Optical Society of America, volume 4, number 1, p.267-280, 1987.

[21] C. Buckberry, J. Davies, "The Application of TV Holography to Engineering Problems in the Automotive Industry", Society for Experimental Mechanics conference proceedings on Hologram Interferometry and Speckle Metrology, p.268-278, 1990.

[22] T. R. Judge, C. Quan, P. J. Bryanston- Cross, "Holographic Deformation Measurements by Fourier Transform Technique with Automatic Phase Unwrapping", submitted to Optical Engineering.

[23] N. Deo, "Graph Theory with Applications to Engineering and Computer Science", Prentice-Hall series in Automatic Computation, 1974.

Chapter 4

Accuracy and Resolution Analysis of Holographic Fringe Data

This chapter presents an investigation into the resolution and accuracy of the measurements which can be obtained with holographic interferometry. The analysis will be aimed at an understanding of the phase measurements attainable. This, in turn, implies the accuracy obtained in the measurement parameter via the sensitivity vector for solid surfaces, or an equivalent statement for other types of fringe pattern.

The measurement resolution is governed by the phase-stepping algorithm used and the hardware which is available. The CCD camera and framestore determine the spatial resolution of the images, and the digitiser in the framestore sets the number of grey levels which represent the intensity of the fringe field.

The accuracy is inherently less than the resolution due to the presence of random and systematic errors. In the single reconstruction beam method presented in the previous chapter, phase errors are deliberately added to the fringe field to facilitate the phase stepping approach. This is the main systematic error present.

In the following section the phase measurement resolution is determined. The factors which can reduce this resolution are then investigated.

4.1 Resolution of the Phase Stepping Algorithms

The resolution of the phase stepping approach is determined by the hardware which is available, and the type of algorithm used. The current commercial specification for a framegrabber/framestore and CCD camera combination is a spatial resolution of 512 x 512 pixels with a grey scale range from 0 to 255, i.e. 8 bits.

The digitisation process applied to the otherwise continuous intensity data implies the resolution which can be attained. This is found by simulating the phase stepping process with computer-generated data. A set of four phase stepped cosinusoidal fringe fields are shown in figure 4.1. The data has been generated to model the case where the full intensity range of the capture device is utilised. A phase step of 90° has been used in this case. The wrapped phase map has then been calculated by applying the four phase step approach described in the last chapter, (see section 3.5, equations 3.39 and 3.40), see figure 4.1. Owing to the noise-free nature of the generated fringes, a conventional phase unwrapping algorithm can be applied successfully giving the distribution in the top graph of figure 4.2. This data is then subtracted from the known phase distribution which was input to the simulation program. The result is shown in the lower graph of figure 4.2, indicating a resolution of $\pm 0.3°$.

The phase resolution of $\pm 0.3°$ can only be achieved in ideal situations. In practice, the actual resolution is limited by the spatial resolution of the camera/framestore combination. The spacing of the fringe data relative to the spatial resolution of the detector implies the number of actual data samples per fringe which are obtained. To achieve a resolution of 0.3° over a detector of width 512 pixels requires the phase to vary by 0.3° per pixel. This implies a minimum fringe period of 1200 pixels. In practice, the average resolution is found from the ratio of the phase change across the detector to the number of pixels in the same direction.

Several other factors affect the resolution of the phase measurements which can be achieved in practice.

i) Fringe modulation

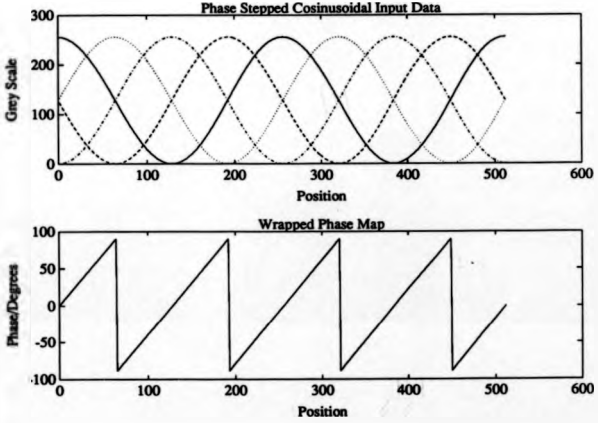


Figure 4.1: Input Data to Phase Stepping Algorithm Simulation

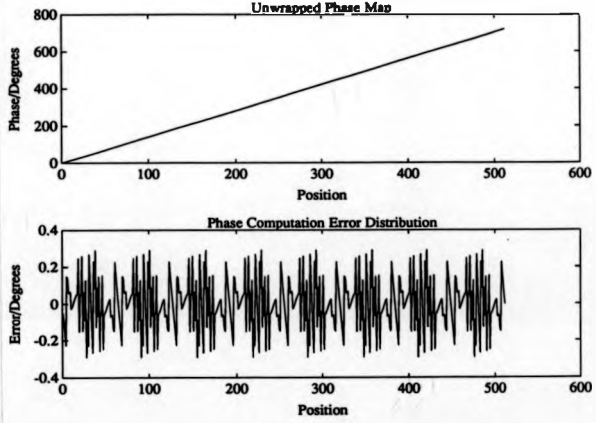


Figure 4.2: Phase Measurement Resolution in Phase Stepping Analysis

ii) The noise level present in the fringe fields.

iii) Inaccuracies in the phase step values from those expected.

iv) The magnitude of the phase step introduced.

v) Non-linearities in the detector transfer function.

vi) The phase stepping algorithm used.

Each of the points above will be examined using five different phase stepping algorithms. The algorithms will be described in turn.

The simplest algorithm which has been suggested uses just two fringe fields, normally with a 90° phase step between them [1]. Using the same notation as equation 3.38, the fringes are described by :

$$\begin{aligned} i_0(x,y) &= i_2(x,y)\cos\phi(x,y), \\ i_{-\pi/2}(x,y) &= i_2(x,y)\sin\phi(x,y). \end{aligned} \tag{4.1}$$

As only two equations are used, the background intensity, i_1 in equation 3.38, must be filtered out prior to the application of this algorithm. The ratio of equations 4.1 eliminates the fringe modulation i_2 giving

$$\phi(x,y) = \tan^{-1}\left(\frac{i_{-\pi/2}}{i_0}\right). \tag{4.2}$$

The second algorithm uses three fringe fields with the same intermediary phase step. The background intensity is also eliminated from the solution for the phase making the algorithm more robust. The phase is calculated by equation 3.39, and gives :

$$\phi = \tan^{-1}\left(\frac{(i_{2\theta}-i_0)(\cos\theta-1)-(i_\theta-i_0)(\cos 2\theta-1)}{\sin\theta\,(i_{2\theta}-i_0)-\sin 2\theta\,(i_\theta-i_0)}\right). \tag{4.3}$$

Thirdly, a further degree of freedom is allowed by using four images with a constant but unknown phase step. This approach is favoured because it removes the need for

absolute calibration of the phase stepping mechanism and it gives a fully automatic method to calculate the wrapped phase map. The phase step is calculated from equation 3.40, giving :

$$\theta = \cos^{-1}\left(\frac{i_0 - i_\theta + i_{2\theta} - i_{3\theta}}{2(i_\theta - i_{2\theta})}\right). \tag{4.4}$$

The average phase step calculated over many image points, together with 3 consecutive images, are then used in equation 3.39 to determine the phase.

The final two algorithms are both applicable when a 90° phase step is used. When four images are captured, the intensities in the fields are given by :

$$\begin{aligned} i_0(x,y) &= i_1(x,y) + i_2(x,y)\cos\phi(x,y), \\ i_{\pi/2}(x,y) &= i_1(x,y) - i_2(x,y)\sin\phi(x,y), \\ i_\pi(x,y) &= i_1(x,y) - i_2(x,y)\cos\phi(x,y), \\ i_{3\pi/2}(x,y) &= i_1(x,y) + i_2(x,y)\sin\phi(x,y). \end{aligned} \tag{4.5}$$

The phase is found from the expression :

$$\phi(x,y) = \tan^{-1}\left(\frac{i_{3\pi/2}(x,y) - i_{\pi/2}(x,y)}{i_0(x,y) - i_\pi(x,y)}\right). \tag{4.6}$$

This introduces four equations to eliminate two variables and calculate a third. Thereby some redundancy is built in to the algorithm. This idea was taken further to use five images with four 90° phase steps [2, 3]. The phase is calculated using [3],

$$\phi(x,y) = \tan^{-1}\left(\frac{2\left(i_{\pi/2} - i_{3\pi/2}\right)}{2i_\pi - i_{2\pi} - i_0}\right). \tag{4.7}$$

4.1.1 Resolution of Phase Stepping Analysis - Dependence on Fringe Modulation

The first parameter investigated is the fringe modulation i_2. The parameter is expressed as a percentage of the maximum amplitude possible in an 8 bit image, i.e. 100% represents an amplitude of 255 and 50% an amplitude of 127 grey scales. The same phase distribution shown in figure 4.1 is used to generate the fringe data and also serves as the reference against which the results are compared. Following Kerr [1], the standard deviation of the calculated phase values against the input phase distribution is used to assess the accuracy of the data. The standard deviation is a measure of departure from the linear phase distribution used as input to the simulations. For the case of reduced fringe modulation, all the phase stepping approaches give identical results. This is expected as all the algorithms allow for a variation in the fringe modulation i_2. The form of the standard deviation measurement error is shown in figure 4.3.

The magnitude of the error does not become significant until the modulation decreases to less than 10%. Even at this point, the standard deviation phase error is less than 2° which is equivalent to $\frac{1}{180}$th of a fringe. In practice, the fringe modulation does not fall below 50%, and the reduction in phase resolution is not significant.

4.1.2 Resolution of Phase Stepping Analysis - Dependence on Image Noise

In the examination of diffusely reflecting surfaces, speckle noise is present in the fringe fields. The characteristics of speckle noise are such that the magnitude of the deviation from the 'true' intensity is larger for higher true intensities. This has led to the multiplicative noise model for speckle effects [4, 1]. Mathematically, the simulations have been performed by calculating the intensity in the phase stepped fringe data $i(x, y)$ from the true intensity $i_t(x, y)$ using the expression

$$i(x, y) = i_t(x, y) + pni_t(x, y), \tag{4.8}$$

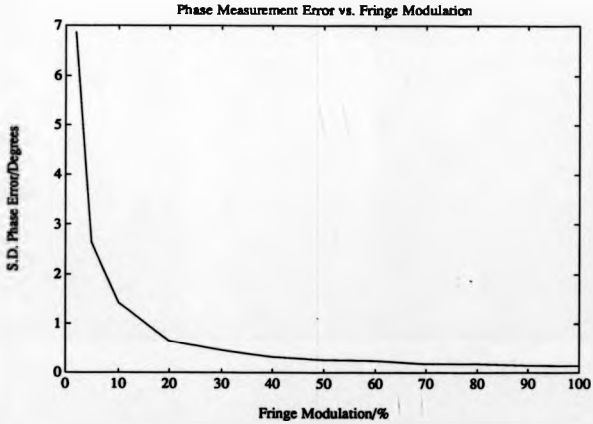


Figure 4.3: Phase Measurement Errors resulting from a Reduction in Fringe Modulation

where p is the percentage noise introduced, and n is a random number in the interval $[-1,+1]$. By variation of p, the noise content in the data is controlled. Simulations were performed for noise levels in the range 0% to 40%. The standard deviation error was calculated with respect to the noise-free phase distribution from the corresponding phase-step algorithm. This merely serves to give a zero error for the case of zero noise. The error distributions for the five algorithms are shown in figure 4.4.

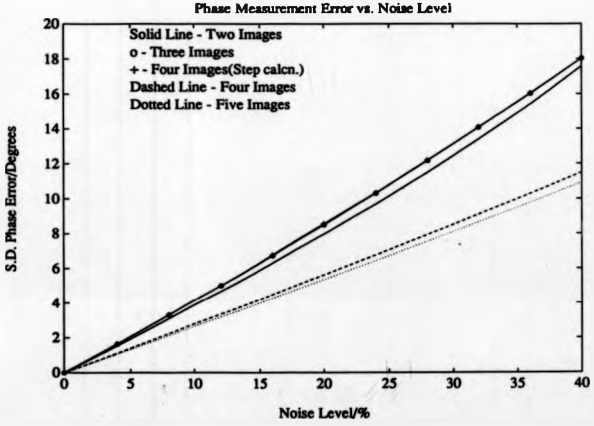


Figure 4.4: Phase Measurement Error Variation vs Image Noise

Similar behaviour is seen for the four-image algorithm (incorporating the phase-step calculation) and the two and three image methods. The variation is approximately linear with 10% noise giving a standard deviation phase error of $< 4°$, i.e. $\frac{1}{80}$th of a fringe. This level of noise has been found in raw holographic fringe data. The graph indicates that to attain a fringe resolution of $\frac{1}{100}$th part, the noise content

in the images must be less than $\approx 6\%$.

For the four and five image algorithms, which incorporate some redundancy into the solution, a similar variation is seen but of reduced magnitude. Noise levels up to $\approx 11\%$ can be tolerated to maintain accuracies of $\frac{1}{100}$th of a fringe.

4.1.3 Resolution of Phase Stepping Analysis - Dependence on Detector Linearity

In all the phase-stepping analysis simulations carried out so far it has been assumed that a perfectly cosinusoidal fringe pattern is to be analysed. The intensity distribution in the fringes has been shown to be cosinusoidal. However, the digitised data is affected by the transfer function of the CCD camera, or image sensor. This is illustrated in figure 4.5 which shows a non-linear camera transfer function and the distortion in the cosinusoidal fringe pattern which is produced.

A set of four phase-stepped fringe fields have been generated using the non-linear response function of figure 4.5. This data was processed using the four image, phase step calculation method to yield the unwrapped phase maps in figure 4.6. The phase error between the known phase distribution and that formed from the distorted data is shown in the lower half of this figure. The error shows a quasi-sinusoidal profile of magnitude $\pm 20°$.

This example is shown with an exaggerated non-linearity in the camera transfer function in order to illustrate the effect. In practice, the linearity given by solid state CCD cameras is better than 2% [5]. A simulation has been performed for a sensor with a 5% non-linearity. The form of the transfer function is similar to that in figure 4.5 with the maximum non-linearity of 5% at an input grey scale of 255. The peak-to-peak error in the unwrapped phase map produced is $\pm 1.5°$. This represents $\frac{1}{240}$th of a fringe.

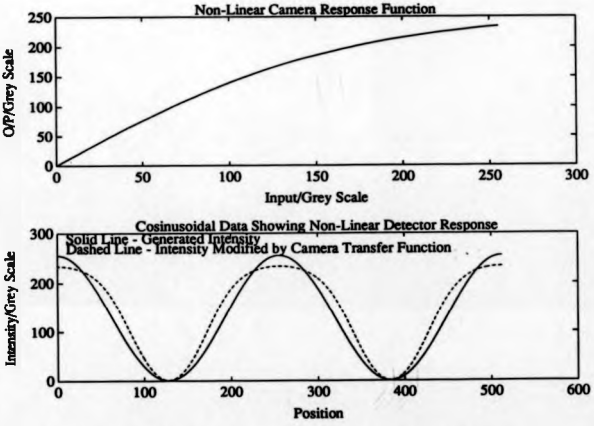


Figure 4.5: Distortion Effects of Non-Linear Camera Transfer Function

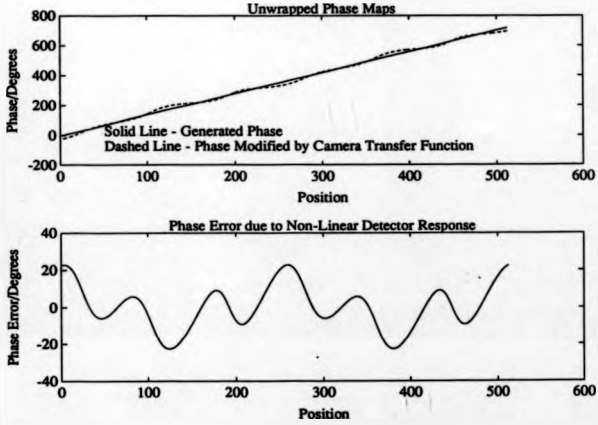


Figure 4.6: Unwrapped Phase Distributions Simulating the Effect of Detector Non-Linearity

4.1.4 Resolution of Phase Stepping Analysis - Dependence on the Phase Step

A major source of error in phase stepping analysis schemes has been shown to be the presence of inaccurate phase steps between the fringe fields [2, 3]. The effect is demonstrated in figure 4.7 which shows the unwrapped phase distributions for zero phase step error and a 20% phase step error. The three image algorithm was used to calculate the phase distribution in this example. The error is seen as an oscillation at twice the frequency of the original cosinusoidal fringes.

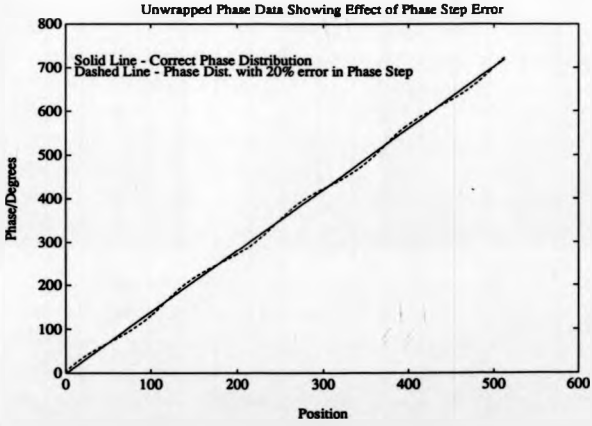


Figure 4.7: Effect of Phase Step Errors on the Unwrapped Phase Distribution

Two types of error distributions can be calculated for phase step errors. Firstly, the difference between the calculated phase and the known phase used as input to the simulation can be found at each point. This represents the absolute error in the phase

measurements. Secondly, the relative error can be determined. This is calculated with respect to a reference position in the fringe data. The phase at the reference position is subtracted from the phase at the point in question for both the calculated phase map and the known phase map. These two values are then compared to give the relative error.

The relevance of the absolute error and relative error parameters depends on how the phase data is to be used. In many cases the relative error is more meaningful as the phase distribution relative to some point in the field is required, for example in contouring an object or measuring a deflection with respect to some point in the field. When the phase data is to be used for strain analysis then the spatial derivatives of the phase are calculated. In this situation the relative phase error determines the error in the strain distribution. The unwrapped phase distributions in figure 4.7 have been offset to show the relative error caused by an error in the phase step between the fringe fields.

The effect of phase step errors on the unwrapped phase map is compared for the five phase stepping algorithms in figure 4.8. This shows the standard deviation of the absolute phase measurement error against an additive error in the phase step from a nominal value of 90°. For the case of a 10% error, the phase steps between the five fields are 99°, 108°, 117° and 126°. The curve for the four-image, phase-step calculation method not surprisingly shows an error of less than 0.5° independent of the magnitude of the phase step error. For the other techniques, the absolute phase measurement error shows an approximately linear variation with the percentage error. As the number of images used is increased, the absolute error is correspondingly increased. This trend has also been presented by Kerr [1]. For the three image algorithm, the standard deviation of the absolute phase measurement error is approximately equal to the magnitude of the phase-step error.

The same calculations have been performed to produce graphs for the relative error, see figure 4.9. Again the four-image, phase-step calculation algorithm produces an error of less than 0.5° for all values of the phase step error. The typical errors produced by the other algorithms show that two images now give the largest errors

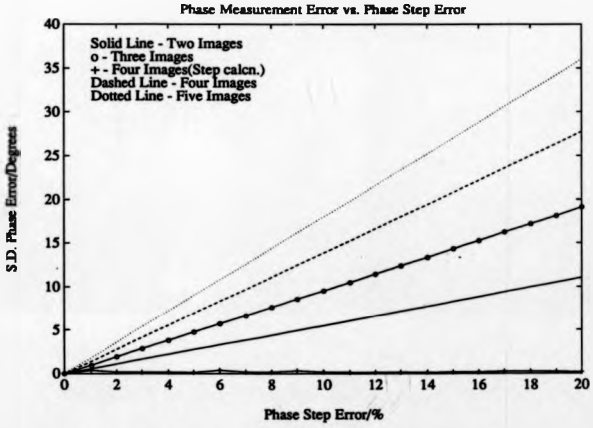


Figure 4.8: Absolute Phase Measurement Errors for Inaccurate Phase Steps

whilst the five image algorithm gives a performance similar to that when the phase step calculation is made. This property of the five image algorithm has previously been shown by Hariharan [3]. By recognising that the result from Kerr's analysis is an absolute measurement error, and Hariharan is calculating the relative error, the apparent contradiction between the two results is resolved.

The overall magnitude of the measurement errors is reduced by a factor of 2 when the relative error is considered rather than the absolute error.

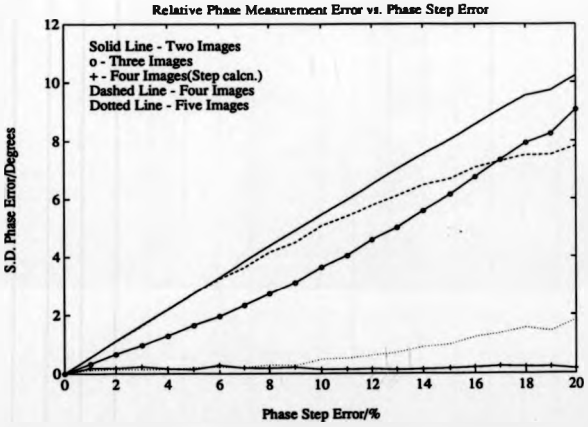


Figure 4.9: Relative Phase Measurement Errors for Inaccurate Phase Steps

The magnitude of the phase step itself also affects the resolution of the phase measurements which can be achieved. This has been simulated using a phase step range from 10° to 175°. The two, four and five image algorithms presented here are specifically designed for use at a phase step of 90°. The four image, phase step

calculation method consists of two stages : calculating the phase step, and then calculating the interference phase using only three images, (see equation 3.39). The interference phase calculation method used here replaced the previous three image algorithm which was also limited to a phase step of 90°.

The results are shown in figure 4.10. The top figure shows the standard deviation of the absolute phase error compared with the known phase distribution. This data therefore includes the error due to the quantisation of the image intensities to 8 bits. The phase stepping approach can be applied with less than 1° error for the phase step range 30° to 165°. The three image method gives a slightly better resolution as the phase step calculation is not required in this case.

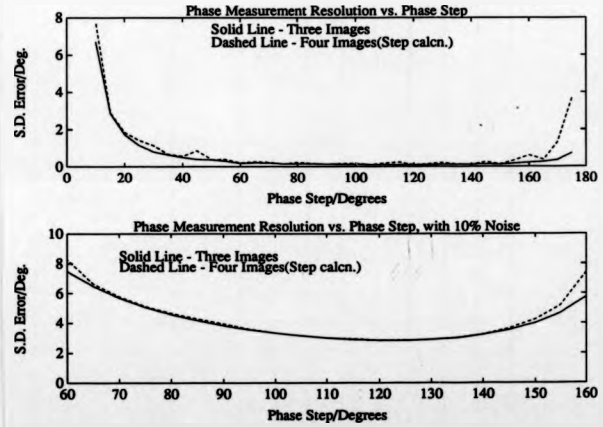


Figure 4.10: Phase Measurement Resolution against Phase Step Magnitude

The dependence of the measurement resolution on the phase step is increased

when speckle noise is added to the fringe data. As described above, raw holographic fringes contain approximately 10% speckle noise. The phase measurement resolution attainable is shown in the lower graph of figure 4.10 for a phase step range of 60 to 160 degrees when 10% multiplicative noise has been added to the fringes. The added noise effectively increases the rate at which the measurement resolution is decreased about the minimum point. The graph shows that a resolution of $\frac{1}{90}$th of a fringe is achieved for a phase step range of 90 to 150 degrees.

4.2 Accuracy Achieved in Phase Stepping Analysis

Two algorithms have been used in this thesis to calculate the interference phase. Firstly, three images with the same phase-step separation have been used. The phase step must be known from calibration of the phase-shifting mechanism. Secondly, four images, also with the same phase-step separation, have been used. In this case the phase step can be determined prior to calculating the interference phase.

The measurement resolutions which can be achieved using these two methods are compared in the following sections.

4.2.1 Phase Calculation with Three Images

The resolution in the three image algorithm is limited by the noise content of the fringe field. Raw holographic fringe data has been found to contain approximately 10% noise. From figure 4.4 this noise level produces a standard deviation phase measurement error of 4 degrees, or $\frac{1}{90}$th of a fringe. The fringe modulation can be reduced to less than 20% of the 256 grey scale range without decreasing this resolution.

The other major error source which has been found from the simulations described above is due to uncertainty in the value of the phase-step between the fringe fields. It has been shown that the standard deviation of the absolute measurement error is

approximately equal to the magnitude of the error in the phase step value. When the phase is measured relative to some point in the field, this error is reduced by a factor of two. Calibration of the phase-shifting mechanism, i.e. the viewing point movement on the linear traverse, has been described in section 3.3. The method used is to correlate two fringe patterns with a 2π phase shift between them. The intensities in the two fringe fields should be the same in this case, with the effect of non-linear detector characteristics cancelling out. The limitation of the method is then dependent on the noise level in the electronics : the CCD camera, and framegrabber. Experimentally, it has been found that repeatable calibrations can be made to $\frac{1}{100}$th of the traverse distance for a 2π phase change. Therefore the standard deviation of the absolute phase measurements is limited to $\frac{1}{100}$th of a fringe owing to this level of uncertainty in the phase step.

The resolvable limit due to inaccurate phase steps is therefore slightly smaller than that expected due to noise in the fringe fields. The two error sources are independent. The variance of the phase measurement is equal to the sum of the variances of the independent variables [6, p.80-81]. Therefore the total standard deviation expected in the phase measurements is $\frac{1}{68}$th of a fringe, which is equivalent to $\approx 5°$. This has been calculated for a phase step of 90°. From the lower graph of figure 4.10 the same resolution would be expected for a phase-step range from 90° to 150°, the optimum phase step being 120°.

4.2.2 Phase Calculation with Four Images

The four-image algorithm has the added versatility of being able to execute automatically. Calibration of the phase shifting device is also not required. From the simulation results, the measurement resolution which can be attained would be expected to be the same as that for the three image method, (see figure 4.4). However, the result for the average phase step becomes biased when the fringe area sampled does not cover a whole number of fringe periods. This effect is increased when non-linearities are present in the image detector. Therefore the measurement resolution using four

images is found to be slightly less than that with the three image algorithm. This can be seen from the lower graph of figure 4.10, and is a function of the phase step magnitude.

The accuracy of the average phase step parameter has been determined experimentally. The magnitude of the phase step error can then be taken as the absolute measurement resolution in the interference phase as for the three image method.

The holographic vibration data shown in figure 1.6 was used to establish the practical resolution of the four-image algorithm. The traverse was calibrated to an accuracy of $\frac{1}{100}$th of a fringe using the chi squared approach. Two sets of four fringe fields were captured with phase steps of 120° and 80°. Three of the fields and the known phase step could be used at this stage to calculate the interference phase to a resolution of 5° as described in the previous section.

The average phase step over 40000 points was calculated using the four-image algorithm. Results of 114.3 and 88.6 degrees were obtained respectively. Plots showing the distribution of the phase step angle are shown in figures 4.11 and 4.12. The curves show discontinuities at multiples of 30 degrees due to an image quantisation effect. These results correspond to an error of $\frac{1}{60}$ and $\frac{1}{40}$ of a fringe respectively based on the calibration data.

The measurement made for the expected phase step of 120° is close to that achievable for the three image algorithm, i.e. 5°. The bias in the average phase step parameter for this value of the phase step is therefore small. This is expected as the four images in this case are equally separated over one complete fringe cycle.

For the expected phase step of 80° the resolution is decreased to $\approx \frac{1}{40}$th of a fringe, or 9°. The resolution would be expected to decrease by $\approx 1°$ when the phase step is changed from 120° to 80° from the lower graph of figure 4.10. This would give a resolution of 6°. The experimental data shows a resolution of 9°, the extra loss of resolution is due to the bias of the average phase step estimate.

This analysis shows that the four image algorithm can be used to give phase measurements to a resolution of $\frac{1}{60}$ of a fringe over a phase step range of 80° to 150°. When the phase step is at the optimum value of 120°, the resolution is increased to

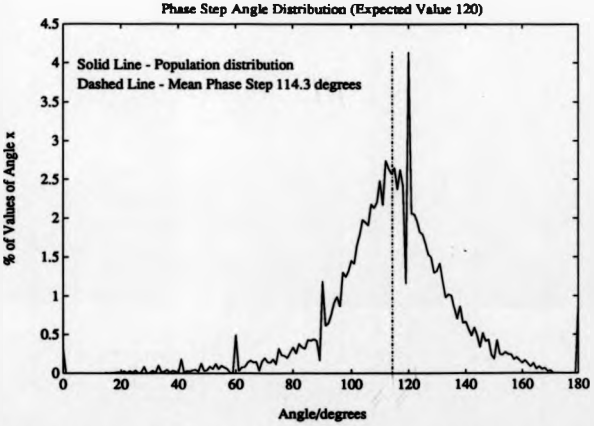


Figure 4.11: Calculated Phase Step Distribution - Expected Value 120 Degrees

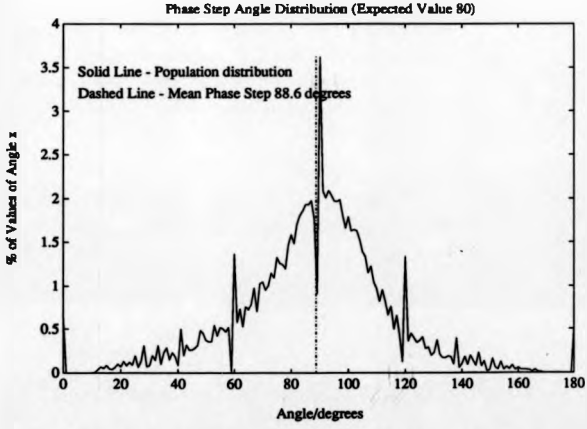


Figure 4.12: Calculated Phase Step Distribution - Expected Value 80 Degrees

$\frac{1}{60}$ of a fringe. This result has been verified experimentally.

4.3 Accuracy Limitations for Single Reconstruction Beam Holographic Interferometry

4.3.1 Holographic Recording and Reconstruction Geometries

To facilitate the error analysis, the geometry of the holographic recording and reconstruction systems must be established. The initial criteria which determines the optical arrangement is the change in sensitivity vector between the image capture positions on the traverse. The sensitivity vector, $\underline{K}$, is defined from section 2.3.1 to be

$$\underline{K} = \underline{k}_1 - \underline{k}_2, \tag{4.9}$$

where $\underline{k}_1$ is the illumination direction, and $\underline{k}_2$ the observation direction. Between the two image capture positions, the observation direction changes from $\underline{k}_2$ to $\underline{k}'_2$, see figure 4.13.

Figure 4.13: Reconstruction Arrangement introducing Notation for Sensitivity Vector Error

From figure 4.13,

$$\begin{aligned}
\underline{k}_1 &= \frac{2\pi}{\lambda}(-\sin\beta\underline{i} - \cos\beta\underline{k}), \\
\underline{k}_2 &= \frac{2\pi}{\lambda}(\underline{k}), \\
\underline{k}_2' &= \frac{2\pi}{\lambda}\left(\frac{\Delta y}{\sqrt{s_O^2 + \Delta y^2}}\underline{i} + \frac{s_O}{\sqrt{s_O^2 + \Delta y^2}}\underline{k}\right),
\end{aligned}$$

where Δy is the distance between the image capture positions, s_O is the object to camera distance, β is the incidence angle of the illumination, and $\underline{i}$, $\underline{k}$ are unit vectors in the x, z directions respectively. The sensitivity vector for the first camera, $\underline{K}$ position is then

$$\begin{aligned}
\underline{K} &= \underline{k}_1 - \underline{k}_2, \\
&= \frac{2\pi}{\lambda}(-\sin\beta\underline{i} + (-1 - \cos\beta)\underline{k}).
\end{aligned} \tag{4.10}$$

After traversing the camera by Δy the new sensitivity vector, $\underline{K}'$ is defined by

$$\begin{aligned}
\underline{K}' &= \underline{k}_1 - \underline{k}_2', \\
&= \frac{2\pi}{\lambda}\left(-\left(\frac{\Delta y}{\sqrt{s_O^2 + \Delta y^2}} + \sin\beta\right)\underline{i} - \left(\frac{s_O}{\sqrt{s_O^2 + \Delta y^2}} + \cos\beta\right)\underline{k}\right)
\end{aligned} \tag{4.11}$$

If the surface displacement at the point P on the object is given by $\underline{L}$, then the phase measurements from the two observation positions are given by

$$\begin{aligned}
\delta &= \underline{K}\,\underline{L}, \\
\delta' &= \underline{K}'\,\underline{L}.
\end{aligned}$$

The phase measurement error is therefore $\delta' - \delta$.

$$\delta' - \delta = (\underline{K}' - \underline{K})\,\underline{L}. \tag{4.12}$$

Substituting equations 4.10 and 4.11 into this expression gives

$$\delta' - \delta = \frac{2\pi}{\lambda}\left(-\frac{\Delta y}{\sqrt{s_O^2 + \Delta y^2}}\underline{i} + \left(1 - \frac{s_O}{\sqrt{s_O^2 + \Delta y^2}}\right)\underline{k}\right)\cdot\underline{L}. \qquad (4.13)$$

In a typical reconstruction arrangement, $s_O = 1500$mm, and $\Delta y = 16$mm is the traverse distance required for a 2π phase change in the fringe field. This gives

$$\delta' - \delta = (-96.6\underline{i} + 0.515\underline{k})\cdot\underline{L}, \qquad (4.14)$$

with $\lambda = 694$nm (i.e. a ruby laser).

A true understanding of the error can be obtained by removing the dependence on $\underline{L}$. This can be obtained with the percentage error, $\delta_{\%}$,

$$\begin{aligned}\delta_{\%} &= \left(\frac{\delta' - \delta}{\delta}\right) * 100, \\ &= \left(\frac{(\underline{K}' - \underline{K})\cdot\underline{L}}{\underline{K}\cdot\underline{L}}\right) * 100, \\ &= \left(\frac{(\underline{K}' - \underline{K})\cdot\hat{\underline{L}}}{\underline{K}\cdot\hat{\underline{L}}}\right) * 100, \end{aligned} \qquad (4.15)$$

where $\hat{\underline{L}}$ is the unit vector (direction) of the displacement.

Two cases need to be considered : when the displacement is out of plane $\hat{\underline{L}} = \underline{k}$, and in plane $\hat{\underline{L}} = \underline{i}$. Taking the more usual, out of plane condition first, the percentage error is,

$$\delta_{\%\perp} = \left(\frac{0.515}{\frac{2\pi}{\lambda}(-1 - \cos\beta)}\right) * 100. \qquad (4.16)$$

The illumination angle normally used to attain out of plane sensitivity can be taken as $\beta = 20°$ (this represents a worse case than normal illumination with $\beta = 0$). The percentage error is then found to be

$$\delta_{\%\perp} = -2.9 * 10^{-3}\%. \qquad (4.17)$$

This is clearly very small and not significant.

The in plane error, $\delta_{\%i}$ is given by

$$\delta_{\%i} = \left(\frac{-96.6}{-\frac{2\pi}{\lambda}\sin\beta}\right) * 100. \tag{4.18}$$

To obtain in plane sensitivity whilst maintaining the current observation direction, the illumination requires rotating to give, for example, $\beta = 60°$. The in plane error is then

$$\delta_{\%i} = 1.2\%. \tag{4.19}$$

This is still not a major limitation. The in-plane error increases as β decreases because the interferometer becomes less sensitive to in-plane motion. With the same illumination direction used for the out of plane calculation ($\beta = 20°$), the in-plane error is

$$\delta_{\%i} = 3.1\%. \tag{4.20}$$

This forms a measurement limit when combinations of in-plane and out-of-plane motions are to be analysed quantitatively. In these situations at least three different sensitivity vectors are required to calculate the vector components of the displacement [7, p.68-77]. One of the arrangements of illumination and observation directions will normally have sensitivity to out-of-plane and in-plane motions. This error then represents a limitation to the accuracy that can be attained. However, in many cases the optical arrangement is made sensitive to only one type of motion, either out-of-plane or in-plane, therefore the error is approximately 1%.

The maximum traverse distance is now defined to be 16mm at an object distance of 1500mm. For the application of the phase-stepping algorithms and traverse calibration, fringe data must be available over a phase step range of 0 to 2π. Therefore a traverse distance of 16mm must produce a phase change of 2π in the fringe field. This criteria defines the spacing of the reference beam fringes in the hologram plane by equation 3.31 which is repeated here,

$$y'_H - y_H = \Delta y\left(\frac{s_O - s_H}{s_O}\right) \tag{4.21}$$

where $y'_H - y_H$ is the distance between the images points on the hologram plane, and s_H is the camera to hologram distance. Typically $s_H = 900$mm, which gives

$$y'_H - y_H = 6.4mm. \tag{4.22}$$

This fringe spacing produces approximately 15 fringes across a 100mm wide holographic plate. The pixel shift in the framestore can also be calculated at this stage using the equation 3.37

$$y'_{IH} - y_{IH} = \frac{s_I \Delta y}{s_O} \tag{4.23}$$

The image distance can be calculated using the thin lens equation and knowing the focal length, f, of the imaging lens which is normally 50mm.

$$\begin{aligned} s_I &= \frac{f s_O}{s_O - f}, \\ &= 51.7mm. \end{aligned} \tag{4.24}$$

Therefore

$$y'_{IH} - y_{IH} = 0.55mm. \tag{4.25}$$

For a CCD camera with a sensor width of 6.4mm containing 512 pixels, this distance corresponds to 44.1 pixels, implying 91.4% of the sensor face is useable.

The fringe separation in the hologram plane of 6.4mm determines the reference beam source positions. Both the case of spherical and collimated beams will be derived.

For spherical beams, the sources are normally off axis at an angle of typically 30 degrees. The geometry is shown in figure 4.14.

The formation of Young's fringes is described mathematically by [8, p.281-283]

$$\Delta = \frac{\lambda l}{b} \tag{4.26}$$

where l is the source to observation plane distance, b is the source separation, and Δ is the fringe separation. Assuming an off axis angle of Θ, the fringe width in the

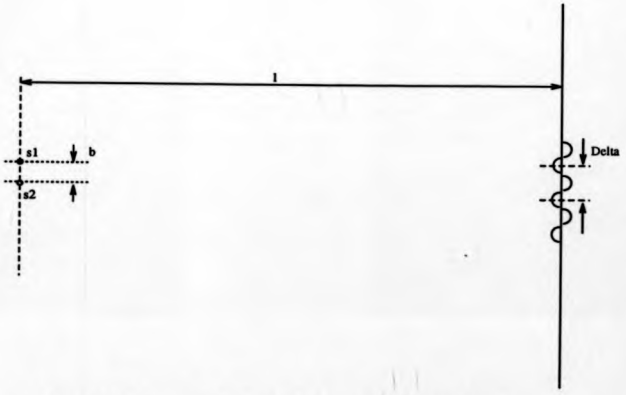


Figure 4.14: Geometry for Spherical Reference Beam Source Positions

hologram plane $\Delta' = y'_H - y_H$ is given by

$$\Delta' = \frac{\lambda l}{b \cos \Theta}. \tag{4.27}$$

This can be re-arranged to give

$$b = \frac{\lambda l}{\Delta' \cos \Theta} \tag{4.28}$$

With $\lambda = 694 * 10^{-6}$, $l = 1000$, $\Delta' = 6.4$, and $\Theta = 30°$ (all distances in mm),

$$b = 125 \mu m. \tag{4.29}$$

Similarly the geometry for collimated beams is shown in figure 4.15.

The fringe separation, Δ is given by

$$\Delta = \frac{\lambda}{2 \sin \theta}, \tag{4.30}$$

with 2θ being the angle between the propagation vectors of the two waves. Again, allowing for off-axis beams at an angle Θ, the fringe separation on the hologram, Δ' is

$$\Delta' = \frac{\lambda}{2 \sin \theta \cos \Theta}. \tag{4.31}$$

With the same definitions as before,

$$\theta = 3.6 * 10^{-3} °. \tag{4.32}$$

The geometry of the recording and reconstruction systems have therefore been defined. The error incurred by the change in sensitivity vector is negligible for out-of-plane displacements. When a combination of out-of-plane and in-plane displacements are present, the error in one of the measurements can be kept to approximately 3%. The corresponding camera traverse distance is 16 mm, and a pixel shift of 44.1 pixels is required to overlay two images with a 2π phase step between them.

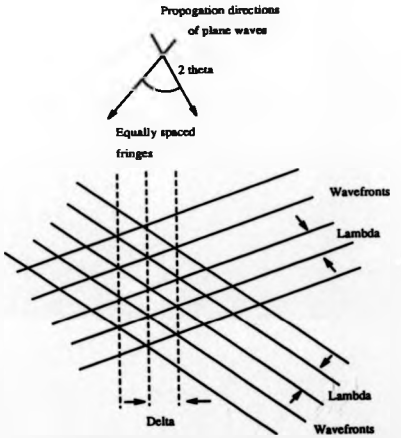


Figure 4.15: Geometry for Collimated Reference Beams

4.3.2 Analysis of the Fringe Pattern Formed using a Single Reconstruction Beam

The effect of reconstructing a dual reference beam hologram with a single beam is to add fringes to the field. Clearly the phase due to the added fringes must be removed in order to calculate the object deformation phase. The phase terms which are present can be examined mathematically by reproducing equation 3.23 for the intensity in the reconstructed fringe field,

$$I_{rec} \propto (2 + 2\cos(\phi_1 - \phi_2 + \psi_{r2} - \psi_{r1})), \tag{4.33}$$

and equation 3.38 used to derive the phase stepping algorithms used

$$i_n(x,y) = i_1(x,y) + i_2(x,y)\cos(\phi(x,y) + \theta_n). \tag{4.34}$$

The object deformation phase $\phi_1 - \phi_2$, equation 4.33, can be represented by ϕ_o, and the phase which is measured ϕ, equation 4.34, can be represented by ϕ_m to avoid confusion. The added phase $\psi_{r2} - \psi_{r1}$ has two effects in the second equation. Firstly, a constant phase is added at each imaged point and represents the reference beam fringe phase ψ_r. Secondly, a phase step can be introduced by changing the viewing position which forms the term θ_n. Therefore $\psi_{r2} - \psi_{r1}$ can be expressed by :

$$\psi_{r2} - \psi_{r1} = \psi_r + \theta_n, \tag{4.35}$$

where ψ_r is a spatially varying function, and θ_n is constant over the image but dependent on viewing position. It follows that

$$\phi_m = \phi_o + \psi_r, \tag{4.36}$$

and the measured phase equals the object deformation phase added to the reference fringe phase.

Two techniques have been investigated in order to compensate for the reference beam fringes such that the object deformation phase can be determined. Both methods work from the unwrapped map of the measured phase.

4.3.2.1 Reference Fringe Compensation - Scheme 1

The form of the reference fringes are expected to be linear across the hologram following equation 3.12,

$$\psi_r(\underline{x}_A) = [(\underline{k}_{r1} - \underline{k}_{r2}) \wedge \underline{w}] \cdot \underline{x}_A + (\Delta\underline{k}_1 - \Delta\underline{k}_2) \cdot \underline{x}_A. \tag{4.37}$$

This suggests calculating the linear part of the unwrapped phase field and subtracting it in order to determine the deformation phase. The phase error as described by the equation also includes any effects due to inaccurate hologram repositioning. Therefore the removal of the linear term will compensate for both the reference beam fringes (the second term of equation 3.12), and the repositioning error of the hologram (the first term of the equation).

For single beam reconstruction, the two beams, $\underline{k}_{p1}$, and $\underline{k}_{p2}$ are equal, therefore

$$\underline{k}_p = \underline{k}_{p1} = \underline{k}_{p2}. \tag{4.38}$$

The difference vectors can then be re-written as,

$$\Delta\underline{k}_1 = \underline{k}_p - \underline{k}_{r1}, \tag{4.39}$$

and

$$\Delta\underline{k}_2 = \underline{k}_p - \underline{k}_{r2}. \tag{4.40}$$

The second term of equation 3.12 is then

$$\begin{aligned} (\Delta\underline{k}_1 - \Delta\underline{k}_2) \cdot \underline{x}_A &= \left(\left(\underline{k}_p - \underline{k}_{r1}\right) - \left(\underline{k}_p - \underline{k}_{r2}\right)\right) \cdot \underline{x}_A. \\ &= -(\underline{k}_{r1} - \underline{k}_{r2}) \cdot \underline{x}_A. \end{aligned} \tag{4.41}$$

The reference fringe phase error is now

$$\psi_r(\underline{x}_h) = [(\underline{k}_{r1} - \underline{k}_{r2}) \wedge \underline{w} - (\underline{k}_{r1} - \underline{k}_{r2})] \cdot \underline{x}_h. \tag{4.42}$$

This equation describes a linear phase term with respect to position on the holographic plate. The magnitude of the repositioning error will typically be much less than that due to the reference beam changes. If the repositioning error is ignored for the time being, the phase error can be described by :

$$\psi_r(\underline{x}_h) = -(\underline{k}_{r1} - \underline{k}_{r2}) \cdot \underline{x}_h. \tag{4.43}$$

The derivation of this equation assumed the plane wave approximation to the phase in the hologram plane. Therefore this equation is only true for collimated beams. The phase error is then linear and can be completely eliminated by subtracting the linear part of the measured unwrapped phase.

When the reference beams are spherical the phase distribution in the hologram plane has small non- linearities, which are normally ignored in the Young's fringe pattern, see equation 4.26. The magnitude of the non-linearity has been determined by a computer simulation which calculates the exact phase distribution in the hologram plane. The reference beam source positions were taken from the geometry described in the previous section, i.e. the sources are 1m from the hologram, at an angle of 30°, and with a separation of $125\mu m$. The range of the calculation was taken over a 125 × 100mm area, representing the standard size of a holographic plate. The phase distribution is shown in figure 4.16.

The main non-linearity present is in the x direction. The magnitude can be found by considering the centre line of the distribution, fitting a straight line, and computing the difference in the results. This is shown graphically in figure 4.17. The top graph shows the original and linear phase distributions, and the bottom graph the difference, or phase error which results after the linear component has been removed. The maximum error is equal to $\approx \frac{1}{8}$th of a fringe.

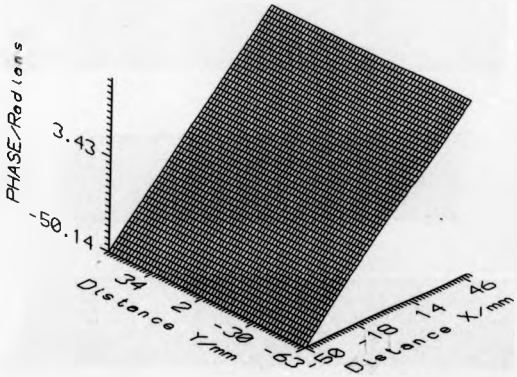


Figure 4.16: Phase Distribution Caused by Dual Reference Beams in the Hologram Plane

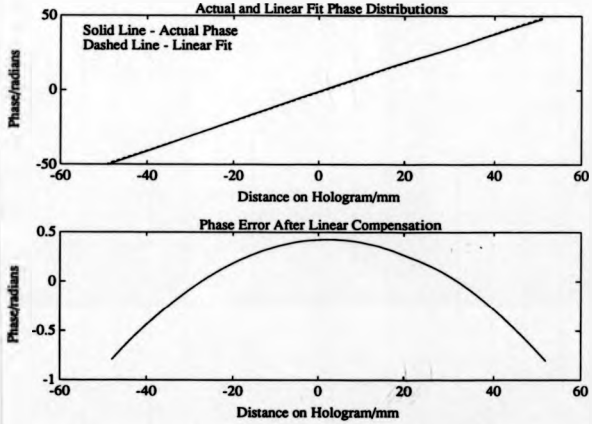


Figure 4.17: Phase Error Remaining after Linear Compensation

A similar computation can be done for the error in the y direction. When the reference beam sources are level with the centre line of the hologram a symmetrical distribution results. Due to the symmetry of the curve there is no gain in fitting a line to the data, and the maximum error is $\frac{1}{60}$th of a fringe. With the sources higher or lower than the hologram centre line the curve is no longer symmetrical and the linear part of the distribution must be removed to reduce the errors to $\frac{1}{60}$th of a fringe.

Simulations were also carried out to show the effect of incorrect hologram repositioning. The sign of the angle deviation determines whether the error after linear compensation is increased or decreased. Even in the case of 10^{o} rotational error about all three axes simultaneously, the error after compensation is less than $\frac{1}{4}$ of a fringe.

Therefore, for spherical beams, subtracting the linear part of the unwrapped phase distribution can be expected to give measurements to an accuracy of $\approx \frac{1}{4}$ of a fringe. This is a considerable reduction from the expected resolution of the phase-stepping method.

A second, practical difficulty, is found in the implementation of this technique regardless of the type of fringes present. It has been assumed that the linear part of the fringe field due to the reference phase can be found. Practically it is not possible to differentiate between a pure rotation of the object and the reference fringes. Hence any set of linear fringes cannot be accurately detected.

This situation is analogous to the problem with the Fourier transform approach where the carrier frequency must be determined in order to translate the spectrum in the frequency plane correctly. Several options are available to determine the reference fringe component.

i) If a part of the object is known to have zero deflection, then this can be used to identify the reference fringes in that region. A similar alternative is to place a static object in the field of view. The fringes across the object then give an accurate assessment of the reference phase error.

ii) In the case of the deformed disc, see figures 3.8 and 3.7, the deformation should be balanced on each side of the central deflection point. The

average slope over an image raster will then have equal contributions of positive and negative slope due to the object deformation. Therefore the average slope provides a good indication of the reference beam phase variation across the hologram. Similar effects are also observed for some vibrations where the same amplitude of vibration is present at neighbouring antinodes.

In the first case some part of an object must be stationary during the experiment, and for the second, knowledge of the form of the deflection must be known.

Assuming that one of these requirements is met, then the accuracy which can be achieved for collimated reference beams is not limited by the error correction procedure. For spherical beams, an accuracy of $\frac{1}{4}$ of a fringe can be expected. These results are still obtained in the presence of hologram repositioning errors.

4.3.2.2 Reference Fringe Compensation - Scheme 2

The removal of the linear component from the fringe field has two main problems : correctly identifying the linear term, and for spherical beams an accuracy of only $\frac{1}{4}$ of a fringe is attainable due to the non- linearities in the Young's fringe pattern. Ideally, a measure of the actual reference fringe field is required. This can be obtained by producing a double exposure hologram with a change in reference beam only, i.e. without deforming the object. For the metal disc, the image obtained from a single beam reconstruction of such a hologram is shown in figure 3.6. This shows the reference fringes with the object in the background. Due to the depth difference between the reference fringes and the object, this fringe pattern can be phase swept by a change in viewing position. The same analysis using three or four phase stepped fringe fields can be performed to calculate an unwrapped phase map. Hence the reference beam phase distribution is found directly. This phase map is then subtracted from the combined object deformation and dual reference beams map to produce the required phase distribution of the deformation.

The accuracy of this approach relies on the repeatability of the reference beam

fringe distribution, and the registration of the two dual reference beam holograms for recording and reconstruction. In practice, registration is achieved by using a well-designed plate holder giving kinematic mounting. Exact repositioning from the recording to reconstruction systems is not needed. The repeatability of the reference beam fringe field is dependent on the overall stability of the recording system, and has not been found to be a problem in practice.

This method is preferred over the first scheme as it removes the need for *a priori* knowledge about the motion of a set of datum points within the field of view. As the method operates by difference, it can be applied equally well to both collimated and spherical reference beams.

4.3.3 Linearity Assumptions for the Imaging Lens

The single reconstruction beam method imposes limitations on the type of lens used to image the fringe pattern on to the CCD. It was found that if the imaging lens could be modelled as a thin lens, then by geometric optics, the relationships between the camera translation and the phase step, and the camera translation with the pixel shift are both linear. The validity of these assumptions are evaluated in the following two sections.

4.3.3.1 Camera Translation to Phase Step Linearity

The relationship between the camera translation and the change in hologram position which the same object point is imaged through is given by 3.31 :

$$y'_H - y_H = \Delta y \left(\frac{s_O - s_H}{s_O} \right), \tag{4.44}$$

for a camera translation of Δy. The distance $y'_H - y_H$ is then related to the phase by knowledge of the reference fringe distribution in the hologram plane. For collimated beams the phase at a point, ϕ_H, is a linear function of position. In one dimension :

$$\phi_H = \phi_O + m y_H \tag{4.45}$$

where ϕ_O is the phase at the origin, and m is the slope of the phase variation. Hence

$$\phi'_H - \phi_H = m\Delta y \left(\frac{s_O - s_H}{s_O}\right). \tag{4.46}$$

Therefore providing the object has negligible depth (s_O constant), the phase step introduced by a given camera translation is constant.

The actual relationship has been determined experimentally in two ways using a reference only fringe image formed with spherical beams. When the beams are spherical the phase variation in the hologram plane is only approximately linear. This effect would be expected to produce a variation in phase step across the reference fringes. A standard 50mm camera lens was used in the experiments.

Firstly, the phase step distribution across an image was calculated. This indicates any deviations present in the imaging system across the field of view. A grid of 10×10 separate regions, each of 20×20 pixels, was examined. The average phase step was computed for each region. The mean of these values was found to be 94.1° with a standard deviation of 4.0°. The data was also examined for definable trends across horizontal rows and vertical columns. For each row or column of 10 values a linear and quadratic curve fit was made and the range of the resulting values computed. The maximum ranges found were ±4.5° for a row, and ±5° for a column. These represent worst case errors of $\approx \frac{1}{72}$ of a fringe (an error in the phase step produces approximately the same size error in the phase, from section 4.1.4). These results for the average phase step in each region are biased according to the number of fringe cycles present in each region. This effect is described in section 4.2.2 and explains the variability found in the results. Therefore the phase-step distribution across an image is a constant to $\frac{1}{72}$ of a fringe.

The phase-step variation expected with spherical beams could show some change across an image row. The phase distribution in the hologram plane due to spherical beams has already been calculated for the first error compensation scheme. From this data, a change in phase step of ±5° would be expected across the whole image. This variation has not been seen in the experimental data because only the central

200 pixels of a 512 pixel row were taken for the calculations, and the object is not imaged over the full width of the hologram surface.

Secondly, the phase step for four sets of four phase-stepped images has been calculated. Each set of images is captured with different separations on the traverse, Δy. The phase step is calculated by equation 3.40, and the results averaged over $\approx$ 20000 pixels to remove some of the effect of speckle and random noise. The same region of the fringe field was used in each case to calculate the phase step. The phase step values tested lie in the range 80° to 120° in order to maintain the accuracy of the calculation. The results are shown in table 4.3.3.1, and plotted in figure 4.18. The figure shows the calculated phase step values and a line of best fit to this data. The errors from the line of best fit to the actual values are also shown in the table, indicating a range of $\pm 0.5°$. Again this is within the expected accuracy of the phase step calculation. Therefore the relationship between the phase step and camera translation is linear.

Image Set	Δy (mm)	Average Phase Step (°)	Error From Linear Fit (°)
1	5.000	116	0.4
2	4.375	104	-0.5
3	3.750	92	-0.2
4	3.125	81	0.3

4.3.3.2 Camera Translation to Pixel Shift Linearity

The relationship between the camera translation and the pixel shift introduced is given by 3.37,

$$y'_{IH} - y_{IH} = \frac{s_I \Delta y}{s_O}, \tag{4.47}$$

where $y'_{IH} - y_{IH}$ is the pixel shift, or translation in the image plane, Δy is the camera translation, and s_I, s_O are the distances from the lens to the image and object planes respectively. This equation is linear for flat objects parallel to the image plane. Again, the expression is verified by experiment, using the same 50mm camera lens.

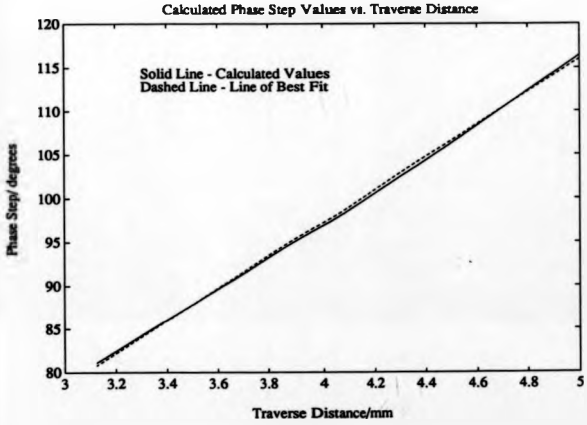


Figure 4.18: Phase Step Variation with Traverse Distance

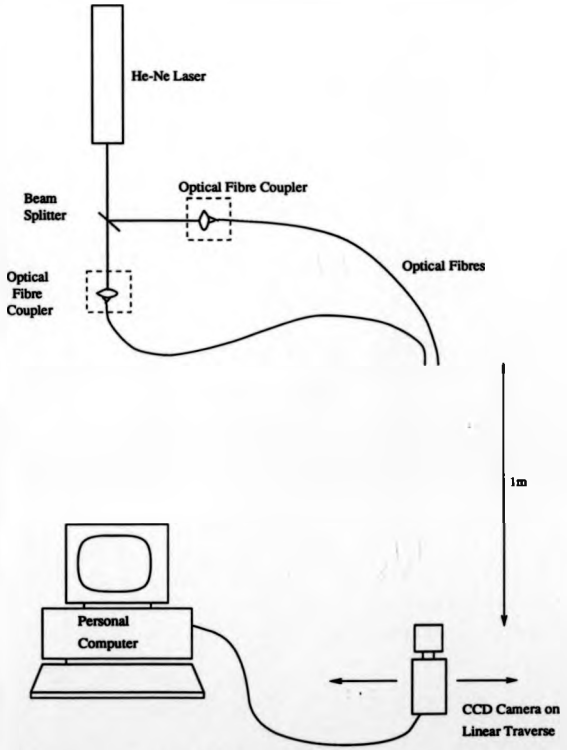


Figure 4.19: Experimental Arrangement to Assess Linearity of Pixel Shifting

A laser beam was split into two parts, each part being fed into a single mode fibre optic by a microscope objective 4.19. The ends of the fibres were used as two point sources which were imaged onto the CCD at a distance of $\approx 1m$. One source was positioned at the edge of the frame and the other towards the centre. The camera was mounted on a linear traverse and eleven images captured at equal traverse displacements of 1mm. The angle subtended by the total translation of 10mm at a distance of 1000mm from the object is equivalent to the example geometry situation of section 4.3.1 where the viewing point was moved by 16mm at an object distance of 1500mm. Each source was correlated with its corresponding position in the other frames, and hence the pixel shift to overlay each of the images was found. After the eleven images were captured, a second image was obtained at the starting position on the traverse. This served as a verification for the stability of the system over the capture time for all the images.

The first and last images captured (both from the same traverse position) were correlated. The results gave a pixel shift of 0.0 for both sources to a resolution of $\frac{1}{5}$th of a pixel. Each of the images from positions 1 to 10 were correlated to the starting position image and the pixel shift calculated. The results are shown in table 4.3.3.2 and plotted in figure 4.20 for both sources. A least squares straight line was fitted to each set of data. The error between the measured pixel shifts and the values from the least squares line are given in table 4.3.3.2. These values show a range of ± 0.15 pixels.

The data shows that there is no difference in pixel shift values across the field of view of the image. The relationship between the pixel shift and the camera translation has been shown experimentally to be linear. Hence the equation shown above is verified.

Δy (mm)	Source 1		Source 2	
	Pixel Shift	Error	Pixel Shift	Error
0.000	0.0	0.07	0.0	-0.15
1.000	4.8	-0.02	5.0	-0.02
2.000	9.8	0.08	10.0	0.10
3.000	14.6	-0.01	14.8	0.03
4.000	19.4	-0.11	19.8	0.15
5.000	24.4	0.00	24.6	0.07
6.000	29.2	-0.09	29.4	0.00
7.000	34.2	0.01	34.2	-0.08
8.000	39.0	-0.08	39.0	-0.15
9.000	44.0	0.02	44.0	-0.03
10.000	49.0	0.13	49.0	0.09

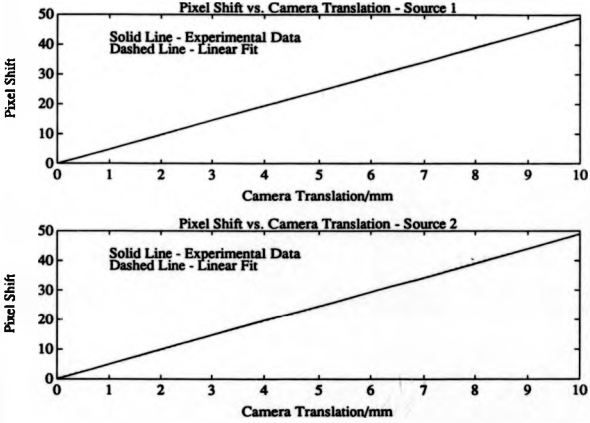


Figure 4.20: Pixel Shift Variation with Traverse Distance

4.4 Analysis of Objects with Substantial Depth

The phase step and image translation equations as stated above exhibit a dependence on the depth s_O between the lens and the object. Variations in the depth of the object will therefore affect the linearity of the relationships between the camera translation and the phase step angle, and the camera translation and the pixel shift. The tolerances of object depth for each of these cases is derived below.

4.4.1 Phase Step Angle Variation with Object Depth

The equation relating the phase step angle variation to object depth and camera traverse distance is repeated below (see equation 4.46).

$$\phi'_H - \phi_H = m\Delta y \left(\frac{s_O - s_H}{s_O}\right), \tag{4.48}$$

where $\theta = \phi'_H - \phi_H$ is the phase step introduced by a camera translation of Δy. The slope of the phase step variation with camera translation m can be calculated from the geometry derived in section 4.3.1. It was shown that a phase step of 360° is applied for a camera translation of 16mm, with $s_O = 1500$mm, and $s_H = 900$mm. Whence $m = 56.25°/mm$.

If θ is the actual phase step required for the mean object depth s_O , then :

$$\theta = m\Delta y \left(\frac{s_O - s_H}{s_O}\right). \tag{4.49}$$

In the case where the actual depth of the object is at s'_O, then the new phase step introduced, θ', is given by :

$$\theta' = m\Delta y \left(\frac{s'_O - s_H}{s'_O}\right). \tag{4.50}$$

The difference $\theta' - \theta$ is then the error in the phase step, θ_{err}.

$$\theta_{err} = m\Delta y \left[\frac{s'_O - s_H}{s'_O} - \frac{s_O - s_H}{s_O} \right]. \tag{4.51}$$

Re-arranging this expression to obtain s'_O gives :

$$s'_O = \frac{s_H}{\frac{s_H}{s_O} - \frac{\theta_{err}}{m\Delta y}}. \tag{4.52}$$

If the allowable phase step error is 3.6° (i.e. $\frac{1}{100}$th of a fringe) and camera translation is 16mm then $s'_O = 1510$mm.

Therefore the allowable object depth about the mean is $\pm$10mm to maintain an accuracy in the phase step to $\frac{1}{100}$th of a fringe across the field of view.

In the case where the object depth variation is larger, the four image, phase step calculation algorithm can still be applied without a loss of accuracy. The algorithm should be implemented such that the calculated phase step at each point is used in the subsequent calculation for the interference phase. This replaces the evaluation of the average phase step over a large area of the image. An alternative algorithm is that proposed by Carre and used extensively by Creath [9]. This approach calculates the interference phase directly from four equi-stepped intensity measurements without knowledge of the phase step.

4.4.2 Pixel Shift Variation with Object Depth

The equation relating the pixel shift variation to object depth and camera traverse distance is repeated below (see equation 3.37).

$$y'_{IH} - y_{IH} = \frac{s_I \Delta y}{s_O} \tag{4.53}$$

where $y'_{IH} - y_{IH}$ is the image translation on the sensor introduced by a camera translation of Δy. The image translation can in turn be related to a number of pixels given the specification of the camera used. In most of the examples shown in this

work a Cohu CCD camera was used giving 512 pixels across a 6.4mm wide sensor [10]. If p is the number of pixels to be translated, then :

$$p = \frac{512 s_I \Delta y}{6.4 s_O}. \tag{4.54}$$

When the object depth is moved to s'_O then the new pixel shift is given by p' :

$$p' = \frac{512 s_I \Delta y}{6.4 s'_O}. \tag{4.55}$$

The error in the pixel shift, p_{err}, is given by $p' - p$:

$$p_{err} = \frac{512 s_I \Delta y}{6.4} \left(\frac{1}{s'_O} - \frac{1}{s_O} \right). \tag{4.56}$$

This expression can be re-expressed to give s'_O :

$$s'_O = \frac{1}{\frac{1}{s_O} + \frac{p_{err} 6.4}{512 s_I \Delta y}}. \tag{4.57}$$

It has been shown in section 3.3 that image translations in the framestore could be resolved to $\frac{1}{4}$ of a pixel. To maintain this level of accuracy a pixel shift error due to object depth must be kept to within $\pm\frac{1}{8}$th of a pixel. For the same geometry as before, i.e. $s_O = 1500$mm, $\Delta y = 16$mm, and a 50mm focal length lens, an object depth variation of ± 4.3mm about the mean value can be tolerated.

4.5 Summary

This chapter has presented an analysis of the accuracies which can be attained using phase stepping analysis methods. Two specific algorithms have been considered : using three images with a known phase step, and four images where the phase step can be calculated. In the first case, the resolution is limited by the noise content normally found in holographic fringe patterns, and the accuracy of the phase step calibration. The standard deviation of the phase measurement error has been calculated from

simulated data to be $\frac{1}{65}$th of a fringe. Secondly, the accuracy of the four image algorithm is limited by the ability to obtain an unbiased estimate of the phase step. From experimental data it has been shown that the phase step can be calculated to $\frac{1}{40}$th of a fringe. This equates to an equal magnitude for the standard deviation of the phase measurement error.

The second half of this chapter considered the particular error sources which are incurred when the single reconstruction beam method is used. In this technique, the fringe pattern caused by using dual reference beams to record the hologram are deliberately added to the field. Two approaches were considered to subtract this data.

Firstly, the phase variation in the reference beam fringe pattern can be assumed to be linear. A linear fit is then made to the unwrapped phase data and this component is subtracted. This procedure is valid for collimated reference beams, but only approximate for spherical beams, where phase errors after linear compensation of up to $\frac{1}{4}$ of a fringe can be expected. The applicability of the method is object dependent in that the deformation of the object may also contain a linear component which cannot be determined correctly.

Secondly, a hologram of the reference beam fringes alone is produced. This provides a direct record of the exact phase distribution to be subtracted. Solution of both the fringe patterns to obtain unwrapped phase distributions can be performed in an identical manner, and the deformation phase obtained. The accuracy of this approach is dependent on the stability of the optical system and the ability to registrate the two holograms.

The geometry of the recording and reconstruction systems are governed by the sensitivity vector change which is obtained when the viewing point is translated. The errors for out of plane measurements have been shown to be negligible. In plane measurements can be made to an accuracy of 1%.

The use of a conventional camera lens to image the reconstructed fringe field has been shown to be valid. Aberrations in the lens would reduce the accuracy of the measurements which can be made. The characteristics of the imaging lens have also

exposed the main limitation of the single reconstruction beam method. This is due to depth variations of the object which cause non-linearities in the phase stepping mechanism and in the pixel shift required to overlay fringe patterns at the same location in the framestore. The accuracy of the method is not reduced providing the object depth varies by less than ±4mm for a mean object depth of 1500mm (measured from the camera lens).

The main limitation of the single reconstruction beam approach is due to the image translation of the fringe data which is required. A resolution of $\frac{1}{4}$ of a pixel has been demonstrated experimentally for the translation process. The accuracy of the phase measurements is then $\frac{1}{4T}$ of a fringe, where T is the local fringe period (in pixels) in the direction of the image translation. To maintain the accuracy of the phase stepping algorithm, $\frac{1}{65}$th of a fringe, the fringe period must be greater than 16.25 pixels.

The standard deviation of the phase errors implies the resolution which can be achieved in surface deformation measurements. For the case where a surface deflection is observed with normal illumination and viewing directions, the accuracy of the deformation measurements for the three image phase stepping algorithm is 5 nanometers; and for the four image, phase step calculation method, the accuracy is 9 nanometers.

Bibliography

[1] D. Kerr, F. Mendoza Santoyo, J. R. Tyrer, "Extraction of Phase Data from Electronic Speckle Pattern Interferometric Fringes using a Single Phase Step Method : a Novel Approach", Optical Society of America A, volume 7, number 5, p.820-826, 1990.

[2] J. Schwider, R. Burow, K. E. Elssner, J. Grzanna, R. Spolaczyk, K. Merkel, "Digital Wave-front Measuring Interferometry : Some Systematic Error Sources", Applied Optics, volume 22, number 21, p.3421-3432, 1983.

[3] P. Hariharan, B. F. Oreb, T. Eiju, "Digital Phase-Shifting Interferometry : A Simple Error-Compensating Phase Calculation Algorithm", Applied Optics, volume 26, number 13, p.2504-2506, 1987.

[4] K. M. Crennell, I. W. Bowler, "The Smoothing of Electronic Speckle Pattern Interferometric Images", Optics and Lasers in Engineering, volume 7, p.163-173, 1987.

[5] C. M. Bastuscheck, "Correction of Video Camera Response using Digital Techniques", Optical Engineering, volume 26, number 12, p.1257-1262, 1987.

[6] K. A. Brownlee, "Statistical Theory and Methodology in Science and Engineering", Wiley series in Applied Statistics, 1965.

[7] C. M. Vest, "Holographic Interferometry", Wiley series in pure and applied optics, 1979.

[8] E. Hecht, A. Zajac, "Optics", Addison-Wesley publishing company, 1974.

[9] K. Creath, "Phase Measurement Interferometry Techniques", in Progress in Optics, ed. E. Wolf, volume 26, p.349-393, Amsterdam, 1988.

[10] Manufacturers specification for COHU model 4710 black and white CCD camera, Cohu, San Diego, California, USA.

Chapter 5

Applications of Dual Reference Beam Holographic Interferometry and ESPI

The application of holographic and interferometric techniques are presented in this chapter. Two distinct sections are considered : the study of amplitude objects, and phase objects.

The fringe fields produced from the examination of amplitude objects are generally solvable using the methods described in the previous chapters. This is because the spatial frequency of the fringes can be resolved within a resolution of 512×512 pixels. As the fringe data is resolvable in a single digital image, holographic and speckle methods are equally applicable. Automatic fringe analysis techniques : filtering, the formation of a wrapped phase map, and phase unwrapping, can then be applied to the whole fringe field.

Holographic deformation data from a series of experiments is presented as examples of the single reconstruction beam method described in chapters 3 and 4.

i) The static deflection of a circular metal disc was used as a laboratory test. A measurement probe placed at a point on the surface monitors the deflection applied. This data was used to verify the accuracy of the

holographic measurements.

ii) An interferogram of the deformed disc was produced with sufficient reference beam fringes to allow the FFT fringe analysis method to be applied. A comparison is made between the FFT results and a phase stepping analysis on the same data.

The experiments were performed with a pulsed laser to record the data, and a continuous wave laser for the reconstruction of the image. The reconstruction errors are the same as for pulsed holography, whether used in vibration analysis, or phase object analysis.

ESPI data, produced using a fibre optic system built at the Rover research centre at Gaydon, is presented. These examples are again of static deformations and illustrate the type of fringe data which is often obtained in an industrial environment. In each case, three fringe fields with phase steps of 90° were analysed. The data provides further examples of the pre-processing method and the operation of the phase unwrapping algorithm.

A study of vibrating objects is made using pulsed holographic interferometry. For this application the source position of the reference beam must be changed between the two laser pulses in order to apply quasi- heterodyne analysis. The fringe pattern formed is interesting as it shows a case where the Nyquist sampling criteria is broken at a particular point in the field. This data illustrates the way in which the phase unwrapping code can accommodate this form of data inconsistency.

The second half of the chapter considers the application of pulsed holographic interferometry to phase objects. A fluid or gas which changes in density causes a change in phase of the light transmitted through the medium. Therefore a fringe pattern can be formed where the fringe orders relate to iso-density contours of the fluid. The examination of phase objects is considerably more difficult than amplitude objects for three main reasons.

i) The fringe data is formed over a larger volume than for a solid sur-

face. This produces large variations in the fringe localisation depth as the fringes tend to localise near regions of large phase gradient.

ii) The number of fringes in the field can range from 10 to several hundred which implies that conventional digital imaging devices have insufficient resolution to sample the whole data field.

iii) One hologram records the integration of the phase changes along the line of sight through the object. In order to provide a quantitative evaluation, at least three such views are required to provide the data to form a tomographic solution for the density field in the gas [1]. This produces a complex experimental arrangement in order to form the three holograms simultaneously.

Automatic fringe analysis of phase objects is an area of current research. The level of development depends on the particular application considered. In this work two examples are given : the examination of high temperature flames, and the study of transonic aerodynamics in a large scale transonic wind tunnel.

Firstly, the examination of high temperature flames is investigated. The recording arrangements for the data are explored and arrangements to record multi-directional data are presented. A simple example of a laminar flame is shown where the single reconstruction beam method has been used to form the unwrapped phase distribution from one direction.

Secondly, a recording arrangement for the study of transonic external aerodynamics has been developed for use in a large scale wind tunnel. The experiment performed gave a flow visualisation only. The results from the experiment are shown, and the possibilities for a quantitative investigation are discussed.

The processing methods used to produce the holographic interferograms are presented in appendix 3.

5.1 The Interferometric Analysis of Surface Deformations

5.1.1 Holographic Analysis of Static Deflections

The holographic system shown in figure 5.1 was used to examine the static deflection of a metal disc. The geometry of the arrangement is similar to that presented in chapter 4, section 4.3.1. A ruby, pulsed laser was used to record the holograms, $\lambda_{ruby} = 694.3nm$. The beam produced by the laser is expanding. A positive lens is positioned close to the laser head in order to produce a collimated beam.

The beam is split into two parts by a prism. The reflection off the front surface of the prism is used as the reference beam and typically contains $\approx 4\%$ of the total beam energy. This beam is expanded by a $\times 4$ microscope objective and spatially filtered. The expanding beam is reflected from a plane mirror on to the holographic plate. The plane mirror is mounted on a linear traverse. By moving the position of the traverse the effective source position of the reference beam is moved, thereby producing the dual reference beams required for hologram recording.

The sample beam is refracted by the prism and then expanded through a negative lens to illuminate the disc. Scattered light from the disc combines with the reference beam to form the hologram. The optics in the sample beam path were coated with a high power threshold dielectric coating to withstand the high peak powers produced by the pulsed laser. A series of single exposure holograms of the disc was then made to optimise the diffraction efficiency and appearance of the holographic image.

The recording of the interferogram was made by two single pulses from the laser. After the holographic plate was placed in its holder the system was allowed to settle for ≈ 5 minutes. Between the two exposures the mirror in the reference beam path can be moved, or the object deformed. This requires a time duration of 2 to 3 minutes. To verify the stability of the system over that time two tests were performed.

i) Firstly, a double exposure hologram with no change in either the object or reference beam was made. The presence of fringes on the object indicates

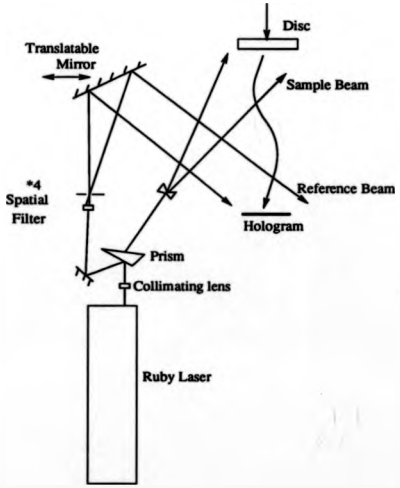


Figure 5.1: Holographic Recording Arrangement to Examine the Deformation of a Metal Disc

an instability in the sample beam path or the way in which the object is mounted.

ii) Secondly, a double exposure hologram of the deformed disc was made to assess the stability of the reference beam path. The appearance of any fringes in the hologram plane indicates an instability.

The geometry of the metal disc is shown in figure 5.2. The disc material was aluminium, with the central deformed region being of 50mm radius and 2mm in thickness. A deformation could be applied at the disc centre by tightening a bolt. At the end of the bolt was a ball bearing to apply the load to the rear surface of the disc. The deformation applied was monitored with a point measurement probe positioned 20mm vertically above the disc centre. The probe used was a Tesa inductive axial movement gauge (IAMG), with a measurement repeatability of 10 nanometers [2]. The gauge was linked to a digital readout system. The stability of the gauge when mounted against the disc was assessed experimentally. The apparatus was allowed to stabilise for over one hour. The output of the gauge was then monitored for $\approx$ 10minutes, showing a variation of $\pm$30nanometers. This variation is principally caused by drift in the electronics of the readout system.

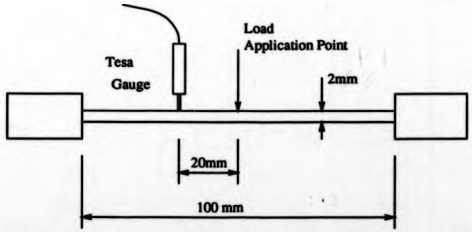


Figure 5.2: Geometry of Disc Object Used for Static Deformation Tests

The holograms were reconstructed using a system similar to that shown in fig-

ure 3.4. A He-Ne laser was used for hologram reconstruction of wavelength $\lambda_{He-Ne} =$ 633 nanometers. The diameter of a He-Ne laser beam is $\approx$ 2mm compared with $\approx$ 12mm for the ruby laser. If both the ruby and He-Ne beams were expanded by the same lens, a different curvature would be produced. Therefore the He-Ne beam diameter was expanded to that of the ruby beam prior to expansion over the holographic plate. This was observed to give better clarity in the image.

The CCD camera was positioned to examine the virtual image formed by the hologram. The camera aperture and laser beam power were varied to optimise the fringe contrast in the fringe field, as described in section 3.3.

Two holograms were used in the subsequent analysis : firstly a dual reference hologram of the deformed disc (the deformation was recorded from the inductance gauge), and secondly, a dual reference hologram of the static object. Both plates were reconstructed using the same system and registrated in the plate holder. By examining the output from the CCD camera on a monitor the approximate traverse distance to give a 2π phase change in the fringe field was found. This distance was interpolated to calculate the distance required to give a phase change of $\approx \frac{\pi}{2}$. Four images were captured with the camera moved by this amount between each frame. The process was repeated for the second hologram using the same traverse positions.

The following analysis stages were then followed.

i) The fringe fields were assessed by the pre-processing algorithm described in section 3.4. The results showed that a single pass of the 3 x 3 pixel averaging filter would increase the area of the fringe field solved.

ii) A region of the fringe field was defined to allow the pixel shift between the images to be calculated. The correlation was performed to $\frac{1}{4}$ pixel resolution and resulted in a shift of 14.5 pixels between each image.

iii) The four image phase stepping algorithm (section 3.5) was applied to the data. The average phase step calculated for the dual reference hologram of the deformed disc was 90°, and that for the reference fringe hologram was 93°. This variation of 3° is less than the figure quoted for the accuracy

of the algorithm of 9°. The difference in the values is due to the random noise present in both sets of data and the change in bias in the average phase step calculation.

iv) Three of the phase stepped images and the average phase step in each case were input to the phase unwrapping algorithm. An appropriate boundary definition for the image was defined to exclude the regions of poor data quality at the edge of the object. The same region from both holograms was solved by the unwrapping algorithm.

A sequence of images showing the results of these operations on the deformation data is given in figures 5.3 5.4 5.5 5.6. These images are : the filtered cosinusoidal fringe data after image translation in the framestore, the wrapped phase map, the unwrapped phase map shown as a normalised grey scale image, and a mesh plot of the unwrapped phase respectively. The effect of the reference beam fringes can be seen as a ramp in the unwrapped phase of figure 5.6.

From the wrapped phase map, figure 5.4, the data appears to contain only a small amount of noise. The unwrapped phase map reflects this in the contiguous distribution which has been produced. The main defect which can be seen is the presence of small 'black holes' in the grey scale image of figure 5.5. These are caused by isolated pixels (or groups of pixels) failing the fringe modulation test for the data at that point. The minimum weight spanning tree approach to solve the phase within each tile exhibits path independency, and is therefore able to unwrap around the low modulation points which are flagged as bad data.

There are no large scale discontinuities in the field. Therefore the solution at the tile level does not represent a severe test of the phase unwrapping strategy. The execution time for all the image analysis routines is $\approx 10 minutes$ on a SUN 4 processor.

The cosinusoidal fringe image and a plot of the unwrapped phase of the reference fringe data are shown in figures 5.7 and 5.8 respectively.

Figure 5.3: Filtered and Translated Cosinusoidal Fringe Pattern of the Deformed Disc

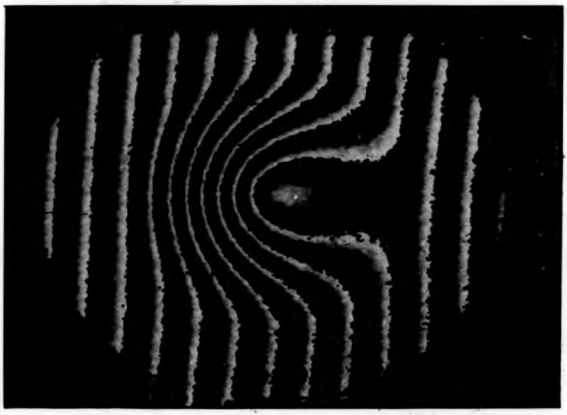


Figure 5.4: Wrapped Phase Map of the Deformed Disc

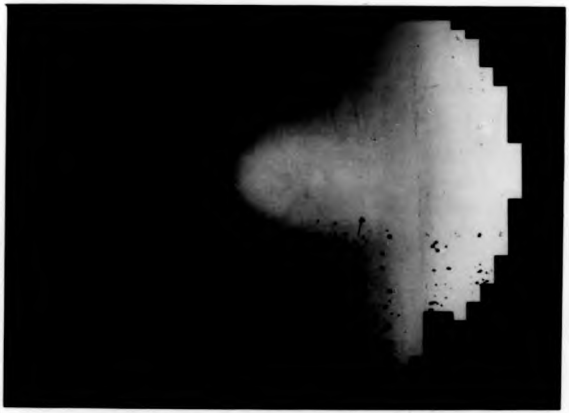


Figure 5.5: Unwrapped Phase Map of the Deformed Disc - Normalised Grey Scale Image

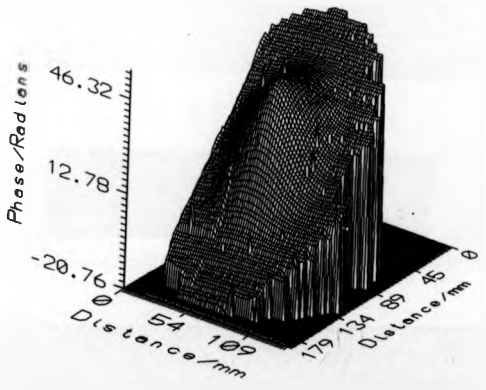


Figure 5.6: Unwrapped Phase Map of the Deformed Disc - Mesh Plot

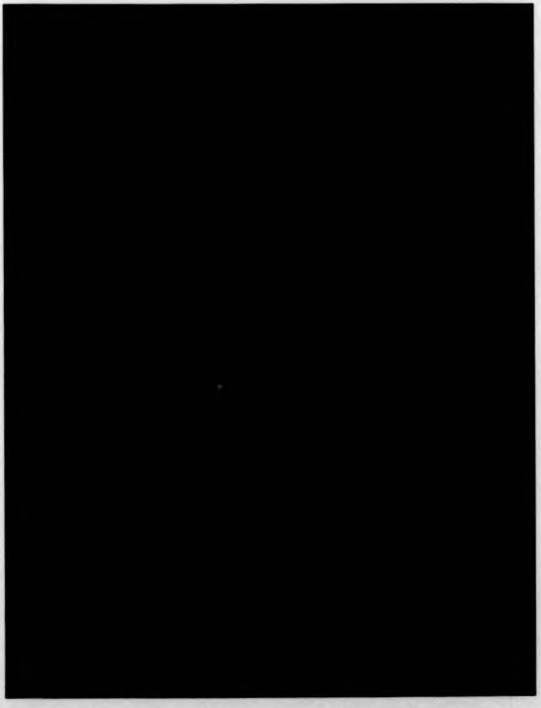


Figure 5.7: Filtered and Translated Cosinusoidal Fringe Pattern of the Reference Beam Fringes

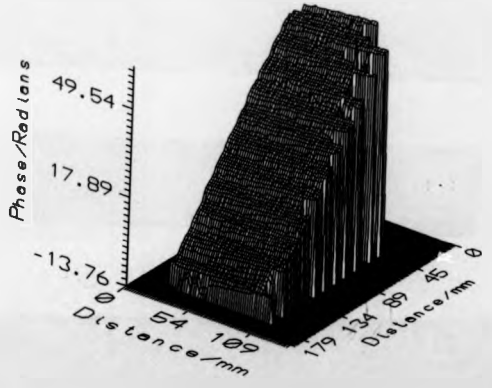


Figure 5.8: Unwrapped Phase Map of the Reference Beam Fringes - Mesh Plot

With the unwrapped phase data in figures 5.6 and 5.8 both error compensation methods described in section 4.3.2 can be evaluated. The phase calculations for each scheme can be verified against the point measurement from the inductance gauge. A deflection of 0.92 microns was recorded.

The sensitivity vector for the recording system is calculated below. The disc is viewed parallel to the direction of the deformation, which is out of plane motion of the disc. The disc was illuminated from an angle of 35° to the normal. Therefore the sensitivity to out of plane motion is given by $\frac{2\pi}{\lambda}(1+\cos 35)$. The phase change, δ, which is expected at the location of the inductance gauge is therefore :

$$\begin{aligned} \delta &= \frac{2\pi}{\lambda}(1+\cos 35).0.92, \\ &= 15.15 radians. \end{aligned} \tag{5.1}$$

This deformation represents 2.41 fringes.

5.1.1.1 Reference Fringe Compensation Scheme 1

In this method a least squares linear fit is made to the phase data in both x and y directions. The fit is calculated over the full area of the unwrapped phase data as the object deformation is symmetrical. The least squares ramp is then subtracted from the phase data. The result of this operation is shown in the right hand plot of figure 5.9, showing a side view of the disc.

As expected, this error compensation scheme is limited by the non-linearities present in the Young's fringe pattern formed by spherical reference beams. The accuracy of the approach can be assessed by examination of the area surrounding the disc. This region should show no deflection. The difference in height between the two ends of the disc is ≈ 0.3 of a fringe. This is larger than that predicted by the simulations in section 4.3.2 due to the reduced distance from the reference beam sources to the holographic plate. This implies an accuracy limit for this compensation scheme when spherical reference beams are used. (If collimated reference beams were used the ramp fringes would be linear and hence this method may be applied without

a loss of accuracy.)

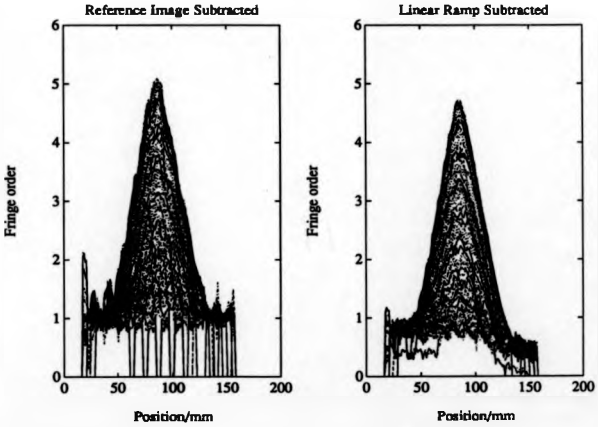


Figure 5.9: Comparison of Reference Fringe Compensation Schemes Applied to Holographic Deformation Data

The graph of the corrected deformation phase also shows an oscillation which is at the frequency of the original fringes. This is caused by a non-linearity in the camera response function. Evidence for this is available from two sources. Firstly, the intensity in the cosinusoidal fringe images is higher on the left hand side of the disc than on the right hand side. Therefore the magnitude of the oscillation in the unwrapped phase would be expected to be higher on the left hand side. This is seen in figure 5.9. Secondly, a cross section of the fringe data reveals an uneven intensity distribution in the fringes. This is shown in figure 5.10 which is the grey scale cross section for a raster of the reference fringe image 5.7. The amplitude of the oscillation is approximately $\frac{1}{20}$th of a fringe.

The other component of the oscillation is caused by the error in the average phase step calculation, performed prior to the formation of the wrapped phase map. The possible error in this case has been shown experimentally to be $\frac{1}{40}$th of a fringe, i.e. half of the observed oscillation.

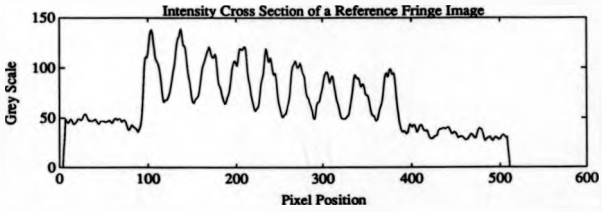


Figure 5.10: Cross Section of Reference Fringe Intensity Showing Distorted Cosinusoidal Fringes

5.1.1.2 Reference Fringe Compensation Scheme 2

The unwrapped phase map due to the reference beams (figure 5.8) is the distribution required for subtraction from the deformation phase in figure 5.6.

Both holograms were recorded at the same traverse positions from the camera, and the holographic plates were registrated in the same position for both recording and reconstruction. Therefore the location of corresponding object points in the

framestore is the same in both images and a direct subtraction of the two unwrapped phase distributions can be performed. This results in the unwrapped phase map shown in the left hand plot of figure 5.9.

The region surrounding the disc is now seen to be at an approximately uniform level. The oscillation in the phase data due to the detector non-linearity is slightly increased for this compensation scheme. This is due to the subtraction of two phase maps both of which contain the error. The amplitude of the oscillation is $\approx \frac{1}{15}$th of a fringe.

The height difference between the disc background and the location of the point measurement gauge was calculated from this data and found to be 2.35 fringes. This shows an error of $\frac{1}{20}$th of a fringe compared with the measured value from the inductance gauge. The gauge resolution of 30 nanometers is equivalent to $\frac{1}{12}$th of a fringe for the sensitivity vector calculated above.

The verification of the measurement accuracy which can be achieved was made difficult due to the presence of a non-linear camera response function. The results show that holographic measurements can be made using this method to an accuracy of better than $\frac{1}{10}$th of a fringe. This represents a conservative estimate and is based on the resolution of the gauge used and an allowance for the non-linearities present in the image data.

The resolution of the method was expected to be $\frac{1}{40}$th of a fringe; limited by the accuracy of the phase step calculation. To verify this figure requires, further experiments with independent measurements of the deformation applied.

Two options are currently being explored to reach this goal. Firstly, the output of the Tesa gauge can be monitored directly without using the digital readout system supplied. Preliminary tests in the nanotechnology group indicate that a resolution of ≈ 5 nanometers can be achieved which is stable over several minutes. Secondly, two components manufactured from a highly stable material such as zerodur could be compared holographically. As the material possesses a very small thermal expansion coefficient, a second measurement could be made some time later using an instrument such as a Talysurf or Nanosurf.

5.1.2 Comparison of the Fourier Transform and Phase Stepping Methods of Fringe Analysis

The holographic reconstruction method which has been described has the effect of adding reference fringes to the deformation data. The hologram which is produced can be analysed by both the Fourier transform method as well as the phase stepping approach providing the fringe pattern is monotonic. This condition can be achieved by using a larger separation of the reference beam sources, thereby giving a greater number of fringes across the holographic plate.

The metal disc experiment described above was repeated using approximately double the number of reference fringes. The hologram was reconstructed using a single beam. Four images were captured from a CCD camera from equi-spaced positions on a linear traverse. One of the pre-processed and translated cosinusoidal fringe images is shown in figure 5.11.

The increased spatial frequency of the fringes has caused the localisation plane to be moved further from the object surface. Consequently, the disc appears slightly defocussed behind the fringes. The fringe data is not ideal for analysis by the Fourier transform method. This can be seen from the cosinusoidal data in figure 5.11. The modulation of the fringes varies considerably across the width of the disc. This is especially noticeable in the fringes about the disc centre. The background, or D.C., level also has variations across the image. As described in section 1.3.5, the variations in the D.C. level, fringe modulation, and deformation phase must all be slowly varying when compared with the carrier frequency. This condition is only partially satisfied in this fringe image.

A further problem arises from the low carrier frequency which has been applied in this case. The period is 13.5 pixels and is only just sufficient to separate the side lobes from the D.C. term of the spectra in frequency space.

These problems are reflected in the appearance of the wrapped phase maps formed by the two methods. The wrapped map formed by the Fourier transform method is shown in figure 5.12; the carrier frequency was subtracted in the frequency domain to

Figure 5.11: Cosinusoidal Fringe Image Suitable for Both Fourier Transform and Phase Stepping Analysis

achieve this result. Aliased fringes are observed in the regions where the carrier was not present in the cosinusoidal fringe pattern. In the phase stepping case, the carrier frequency is still present in the wrapped map, see figure 5.13.

The reduction in fringe modulation around the disc centre causes the noisy data in this region of the wrapped map with the FFT approach. In contrast, the phase stepping method is much less sensitive to this problem (see section 4.1.1) and the data in the same region is of much higher quality.

In both cases the phase was unwrapped using the minimum weight spanning tree (MST) approach. A region was defined in the centre of the image to exclude the surrounding areas of the disc showing aliased fringes and background noise. The Fourier transform wrapped map shows discontinuities in the fringe data near the disc centre. This serves as an example of the phase unwrapping algorithm's ability to isolate areas of bad data. The solution of each tile is performed using the MST at the pixel level. The weighting functions for each edge of the tree linking the tiles include factors for the spatial frequency of the fringe data and the fit of the data between the two neighbouring tiles along the tile edge. The presence of inconsistencies in the fringe pattern will cause the fit, or correlation, between neighbouring tiles to decrease. Therefore the tree is started in the regions near the edge of the disc where the spatial frequency of the fringes is low. Tiles are connected around the disc edge and the centre of the disc is added to the tree from the right hand side where the data is unambiguous. The last tiles added to the tree are those containing the ambiguous data to the left of the disc centre. The results in this area show a discontinuity with neighbouring tiles, as the data cannot be matched at both sides of the tile. The final unwrapped phase is shown as a normalised grey scale image in figure 5.14.

In the case of the phase stepping approach, the wrapped map is comparatively free from inconsistencies. Therefore the data is unwrapped without difficulty. The unwrapped phase data has been processed further to subtract the carrier frequency from the data. This has been performed using the known carrier frequency from the FFT analysis. The result is shown in figure 5.15.

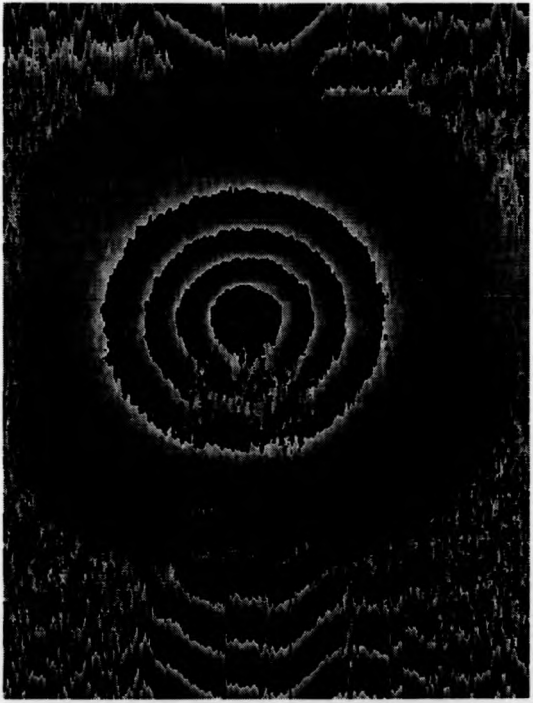


Figure 5.12: Wrapped Phase Map formed by FFT Method

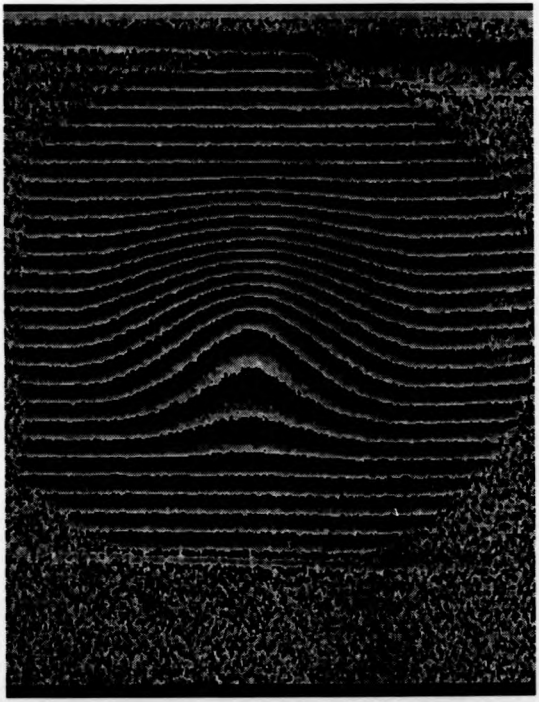


Figure 5.13: Wrapped Phase Map formed by the Phase Stepping Method

Figure 5.14: Unwrapped Phase Map from the Fourier Transform Method - Normalised Grey Scale Image

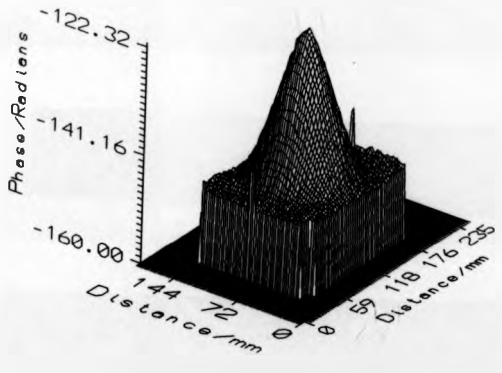


Figure 5.15: Unwrapped Phase Map from the Phase Stepping Method - Mesh Plot

The unwrapped phase data calculated by the two methods has been compared by subtracting one distribution from the other, pixel by pixel. This operation produces the distribution shown in figure 5.16 as a mesh plot. The areas where discontinuities are present in the unwrapped phase from the Fourier transform approach, have been excluded from the calculation. Some large differences are seen at the edges of this plot due to the different location of noise points in the two phase maps.

The standard deviation of the phase distribution has been found to be 21°. The points at the edge of the phase map were excluded from the calculation. This statistic represents the difference between the two solutions, and is equivalent to $\frac{1}{16}$th of a fringe.

This analysis implies that both the phase stepping and Fourier transform methods of fringe analysis can be applied to produce phase maps with a deviation of $\frac{1}{16}$th of a fringe between them. The Fourier transform method would be expected to be less accurate than the phase stepping approach as it uses four times less data as input. A better correlation between the two approaches is thought possible, as the input data was not ideal for analysis using the Fourier transform approach. The use of non-ideal data has shown that the presence of intensity variations across the image can be accommodated with the phase stepping method; whereas significant problems are encountered using the Fourier transform methods. Therefore greater care is required in the application of Fourier transform techniques in order to produce contiguous unwrapped phase distributions of corresponding accuracy to those achieved with phase stepping.

5.1.3 ESPI Analysis of Small Structures in the Automotive Industry

Two examples describing the processing of ESPI data are presented in this section. The images demonstrate the fringe analysis procedures of automatic filtering, formation of a wrapped phase map, and phase unwrapping.

The speckle pattern fringe images were produced by a fibre optic ESPI system,

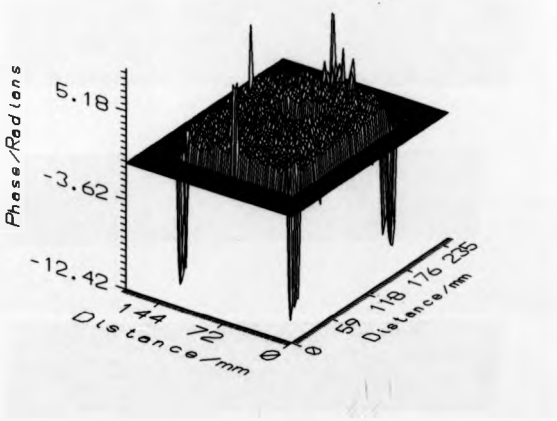


Figure 5.16: Phase Distribution Showing the Difference between the Results from the Phase Stepping and Fourier Transform Techniques

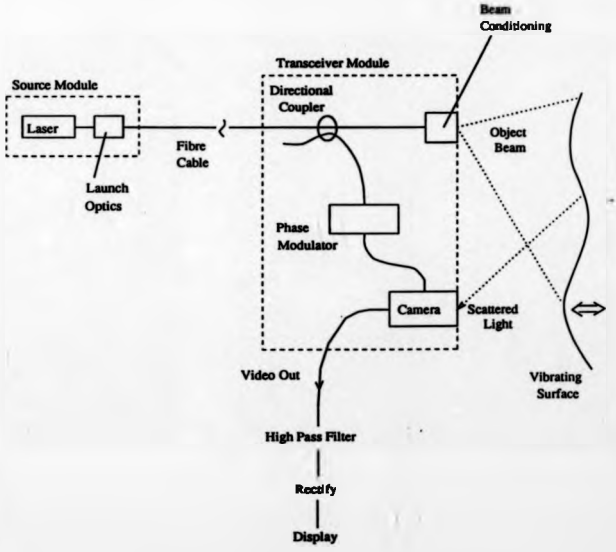
see figure 5.17. This device was constructed by the Rover Group Research Centre at Gaydon, Warwickshire, and a full description is contained in reference [3]. The fringe fields shown here are of static deformations and hence have the same cosinusoidal form as the double exposure holographic data shown above.

An argon ion laser provides the coherent light source which is launched into a fibre optic cable. A reference beam is split off by a directional coupler. The reference beam and scattered light from the object beam are combined at a CCD camera to form the interference pattern. The video signal is high pass filtered and rectified before entering a P.C. framegrabber (of resolution $512 \times 512 \times 8$bits), and high resolution graphics monitor. Phase stepping was achieved by wrapping the reference beam fibre around a piezo electric transducer. The phase stepping mechanism was calibrated by the comparison of straight line fringe patterns with a phase step of 2π.

The two sets of ESPI images examined were, of a cylinder bore, and the cylinder head of a car engine. In both cases an initial frame is captured from the CCD camera of the undeformed object. The torque of one or more of the bolts around the cylinder under examination is then changed. Three frames are captured of the deformed object with different phases of the reference beam, 0°, 90°, and 180°. The fringe patterns were formed by speckle subtraction, i.e. the initial frame was subtracted from each of the frames of the deformed object. This results in three images with phase steps of 90° between them. The fringe patterns represent the change in loading of the cylinder.

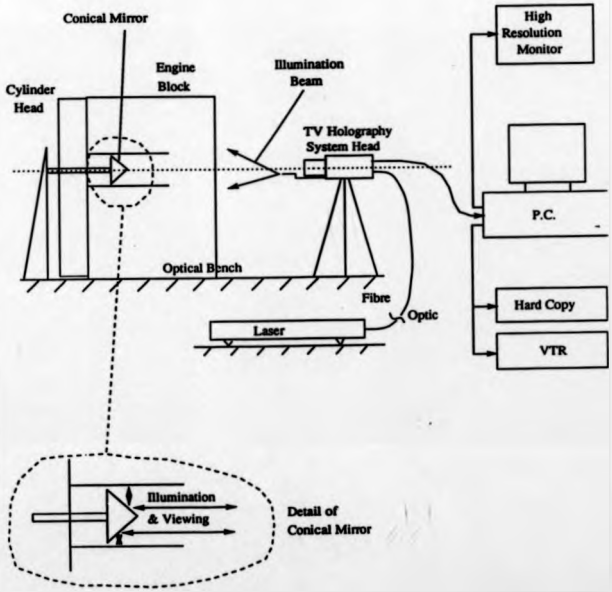
The experimental arrangement for the cylinder bore is shown in figure 5.18. A conical mirror was mounted into the top of a cylinder. When this is viewed from the bottom of the engine, the line of sight reflects off the conical mirror and onto the wall of the bore. Therefore the fringe patterns represent the out of plane deformation of the cylinder walls. The chamber image was produced directly with the camera looking along the bore towards the cylinder head.

The analysis sequence for this data was : pre-processing of the cosinusoidal fringe patterns, formation of a wrapped phase map, and phase unwrapping. The



Fibre Optic ESPI System
and video signal conditioning

Figure 5.17: Fibre Optic Based ESPI System



Detail of Optical System
Applied to Car Engine Block

Figure 5.18: Optical System Used to Examine Car Engine Block

pre-processing of the data has been described in section 3.4. The figures shown in that section were calculated from the bore images. Similar results were achieved for the chamber. The optimum number of passes of the 3 × 3 pixel filter was automatically determined, and four iterations of the filter were performed for each image. The three image phase stepping algorithm was implemented with the known phase step of 90° to form the wrapped phase map. This data was then processed by the phase unwrapping algorithm.

The form of the normalised input images, after preparation are shown in figures 5.19 and 5.23. These are respectively the cylinder bore, and the cylinder head chamber. The wrapped phase maps for the above images are shown in figures 5.20 and 5.24. The data representations used in the two wrapped maps are different : two phase fringes for each cosine fringe are shown for the bore, and a single phase fringe for each cosine fringe for the chamber.

Normalised grey scale images representing the unwrapped phase map for the two cases are given in figures 5.21 and 5.25. In each case the output is unsmoothed. Phase unwrapping within each tile was performed using the minimum weight spanning tree approach for the chamber. A conventional fringe counting method was used to solve the tiles in the bore image. The tile size has been selected as 20 × 20 pixels for both sets of images. Three dimensional mesh plots of the same data are given in figures 5.22 and 5.26. The data is unsmoothed for the chamber, and spike noise has been filtered out of the bore solution.

Finally two figures are given which show the progression of the tree linking the tiles in each case. Figure 5.27 shows the tree for the bore image, and figure 5.28 the tree for the chamber. Each box in the tables represents a tile (the faint regular grid lines). The thick black lines represent fringe wrap-over points, and the thin black line linking the tiles together represents the progression of the MST. Tiles which contain a large proportion of low modulation points are not joined into the MST.

The unwrapped solutions produced show contiguity across objects which are themselves continuous. However, the original phase stepped fringe fields contain regions which are not consistent with each other, as well as pixel scale noise. The effect of

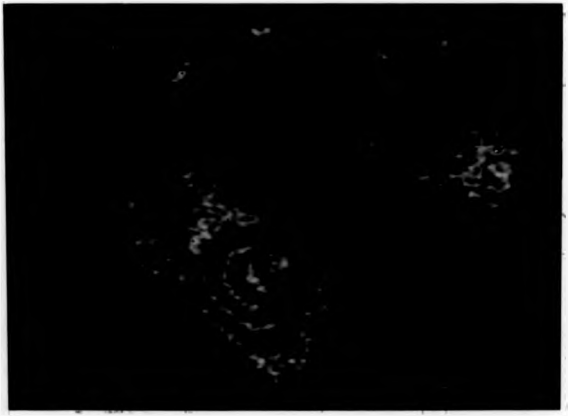


Figure 5.19: Filtered Cosinusoidal Fringe Image of the Cylinder Bore

Figure 5.20: Wrapped Phase Map for the Cylinder Bore

Figure 5.21: Unwrapped Phase Map for the Cylinder Bore - Normalised Grey Scale Image

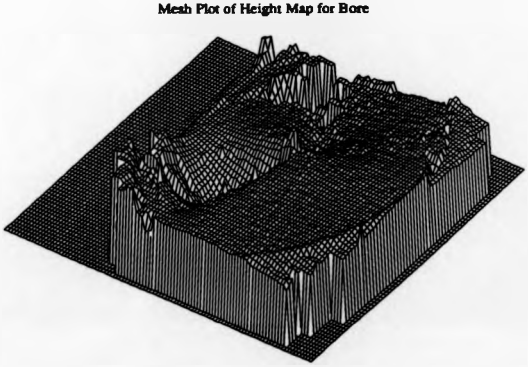


Figure 5.22: Unwrapped Phase Map for the Cylinder Bore - Mesh Plot

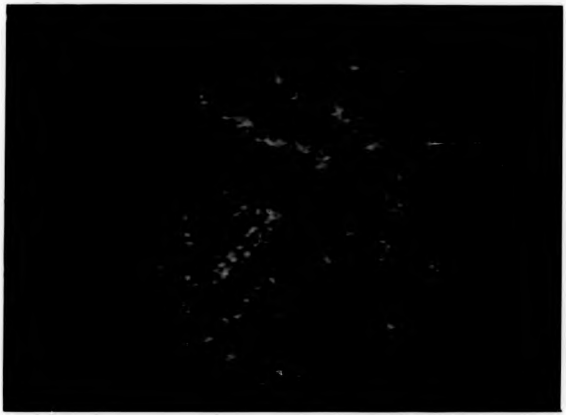


Figure 5.23: Filtered Cosinusoidal Fringe Image of the Cylinder Head Chamber

Figure 5.24: Wrapped Phase Map for the Cylinder Head Chamber

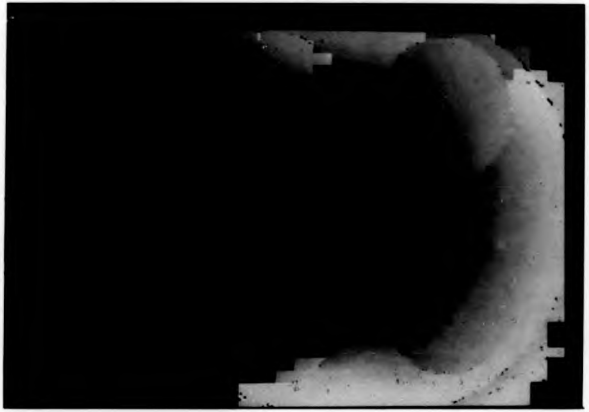


Figure 5.25: Unwrapped Phase Map for the Cylinder Head Chamber - Normalised Grey Scale Image

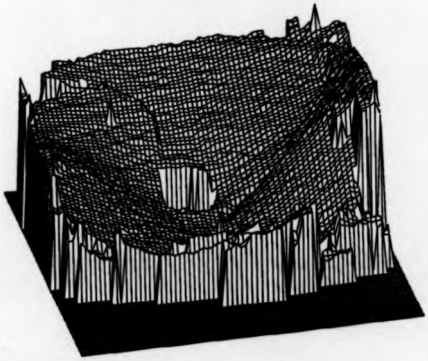


Figure 5.26: Unwrapped Phase Map for the Cylinder Head Chamber - Mesh Plot

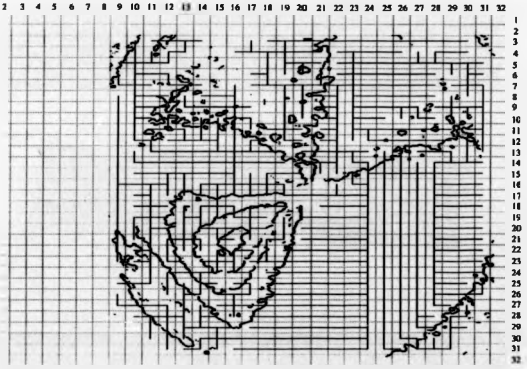


Figure 5.27: Progression of the Minimum Weight Spanning Tree in the Bore Image

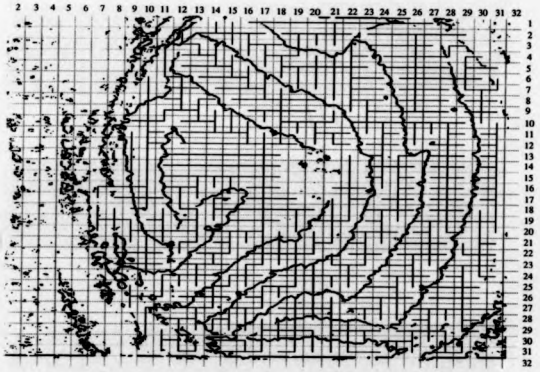


Figure 5.28: Progression of the Minimum Weight Spanning Tree in the Chamber Image

large scale inconsistencies is to cause a step in the unwrapped phase map. This is normally located at a tile boundary. The step occurs when the count for the number of fringe wrapover points between two tiles is dependent on the path taken between the tiles. The occurrence of this type of problem can be traced back to regions of the input data containing aliasing type discontinuities. The structure of the solution through the use of the weighting functions between tiles will minimise the effect of the discontinuities on the rest of the field. This is described with reference to each of the ESPI images below.

5.1.3.1 Phase Unwrapping of the Cylinder Bore Image

The fringe field for the bore (figure 5.20) shows a very marked trough just below and left of the image centre. Around the trough, are fringes of gradually increasing height. Due to the higher spatial frequency of the fringes in this region the tile weights will be higher than for the tiles at other parts of the fringe field. Therefore the rest of the image will be solved first, and the tree will only link the tiles in this region at the ends of the branches. This effect on the development of the tree can be seen in figure 5.27. The depression in the phase map is entered from both sides, i.e. from the regions of approximately uniform phase. These two branches of the tree are not joined across the trough, thereby isolating any discontinuity in the fringe data in that region.

Close to the centre of the field the fringes become very badly disrupted. This is because the fringes were recorded from a conical mirror. At the centre of the cone a large physical distance is seen off a relatively small circular distance. Hence the fringes around the centre are blurred. The outcome may be seen in the grey scale plot where tiles near the centre are unrelated to their neighbours. In the same way the tree is formed around the region of bad data and no connection is made between the tiles on opposite sides (figure 5.27). Here the weighting factors are large due to the low correlation between the tile edges in the region.

The fringe counting method to unwrap the tiles produces striations in the solution,

see figure 5.21. These are caused by a missed fringe edge at a single pixel which then affects the rest of the raster until the tile edge is reached. This effect produces larger discontinuities when larger tiles are used. The resulting spikes and ridges have been removed using a post-processing operation, see figure 5.22. The unwrapped phase at a point is compared with the average phase of its eight nearest neighbours. If the difference is greater than π then the phase at the point is set to the average over the eight neighbours. This process filters out most of the noise spikes from the unwrapped data.

5.1.3.2 Phase Unwrapping of the Cylinder Chamber Image

The wrapped field for the chamber is shown in figure 5.24. The field covers five physically separate objects : the cylinder head itself, and four circular valves. The boundaries between the object surfaces would be expected to give rise to large scale discontinuities in the field in a similar manner to aliasing of the fringe data. The deformation which has been applied was not sufficient to form discontinuities in the fringes between the separate objects towards the centre of the cylinder. The valve edges can be located as a change in direction of the fringes. On the outside edges of the valves, the discontinuities are more apparent. From the solution, figure 5.26, two discontinuous regions are seen as steps in the unwrapped phase near the edge of the fringe data. Note that the mesh plot is rotated by 90 degrees from both the grey scale phase distribution, and the figure showing the progression of the tree. The bottom of the grey scale image is located at the right hand side of the mesh plot.

Considering the two tiles 20,24 and 20,31 in figure 5.28 (referring to x,y tile co-ordinates). The shortest path between them crosses three fringe edges. However the path via tile 15,30 crosses a single fringe edge. Hence the data is inconsistent. The tree will resist crossing a fringe edge until most tiles neighbouring the edge have been solved. This is due to the fringe spatial frequency factor in the edge weight for a tile. The edge is crossed at the point where the spatial frequency of the fringes is least, and the fit along the tile boundary greatest. This occurs at tile 20,26; and subsequently

the height at tile 20,31 is correctly assigned (see right hand side of figure 5.26). The data inconsistency is contained between tiles 15,29 and 15,30 due to the poor data quality in this region.

A similar argument can be applied between tiles 13,14 and 10,14. In this case the inconsistency is contained by the fringe spatial frequency factor in the edge weight.

In the centre of the chamber field there is a hole to locate the spark plug. The tiles in this region of the field have failed because they contain a large number of low modulation points. The absence of any preferential direction in the unwrapping path implies that the tree has been formed around the hole.

The MST method for noise immune phase unwrapping at the pixel level is seen to produce a better, more continuous result than fringe counting. By using the MST, the effect of a noise point is isolated to a single pixel; compare figure 5.25 analysed using the MST, and figure 5.21 analysed using the fringe counting method.

5.1.4 A Pulsed Holographic System for Quantitative Vibration Analysis

The application of double exposure holography to vibration analysis requires the use of pulsed lasers. The pulse duration of a ruby laser is 30 nanoseconds. Over this duration the state of a surface is effectively stationary. Therefore the fringe patterns formed in double pulsed vibration analysis are the same as those for static deformations, i.e. cosinusoidal.

The application of dual reference beam holographic interferometry to vibration analysis requires that the reference beam be changed between the two laser exposures. To achieve this, an optical switching system was used consisting of high speed optical devices, an electronic control system, and a software interface on a personal computer. The system was developed jointly with other members of the department.

The two reference beams are formed by the optical system in figure 5.29. The beam from the ruby laser is collimated and then split 90/10 to produce sample and

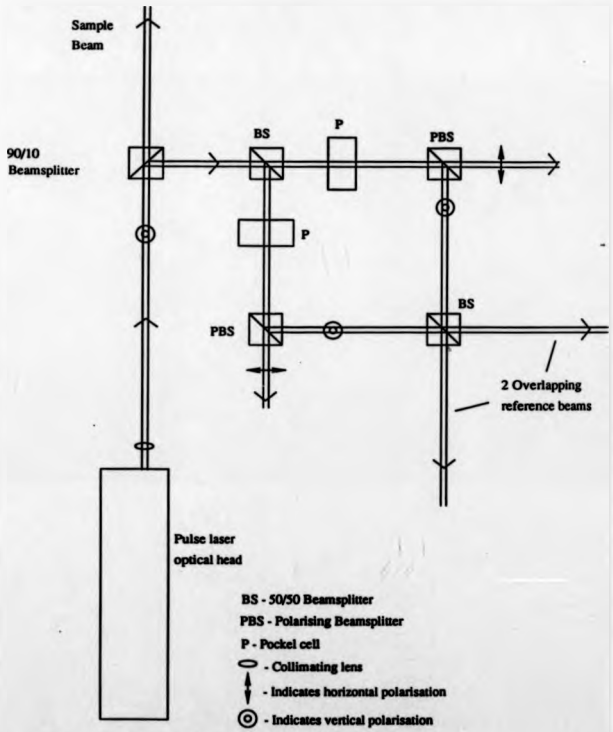


Figure 5.29: High Speed Reference Beam Switching Components

reference beams respectively. The reference beam is in turn split by a 50/50 beamsplitter, each half proceeding to a Pockels cell and a polarising cube beamsplitter before being re-combined at a second 50/50 beamsplitter. The four beamsplitters are located at the corners of a square. By applying a voltage to one of the Pockels cells the beams polarisation is rotated by 90 degrees. Initially the polarisation is vertical. With the Pockels cell off the polarisation is unchanged and the beam is reflected by the polarising cube and reaches the output of the second 50/50 beamsplitter. If the Pockels cell is on the polarisation is rotated to horizontal and the beam is transmitted by the polarising cube and dumped. This produces a switching of the beam at the second 50/50 beamsplitter.

To record the hologram a Pockels cell is switched after the laser pulse is fired, effectively opening a window for the beam. For the second pulse the other Pockels cell is switched. Electronic hardware connected to a P.C. via an RS 232 link takes full control of the pulse laser and external Pockels cell drivers. The hardware incorporates an independent processor which programs timing circuits for the laser flash lamps and each of the Pockels cells, both internal to the laser and external in the reference beam switching system. A software package on the P.C. allows the user to set up the firing parameters which are transmitted to the independent processor. When a 'fire' command is given from software, each timer runs for the preset time and the corresponding action is then performed. The control system has a jitter of less than 10 ns and a resolution of 125 ns. The lower limit for pulse separation from the laser is 0.3 microseconds when using the external Pockels cells.

The external Pockels cell drivers are set to change state for 100 ns. If the Pockels cell is switched with the full voltage required to change its state the polarisation produced will oscillate from vertical to horizontal. In practice, 60% of the full voltage is used to produce a cleanly switched output. The Pockels cell alignment is critical for the overall efficiency of the switch. This can be measured by the intensity of the horizontal and vertical components of the beam which are separated by the polarising cube beamsplitter. An efficiency of better than 99% is achieved in practice.

The apparatus can be used for the quantitative analysis of any transient events

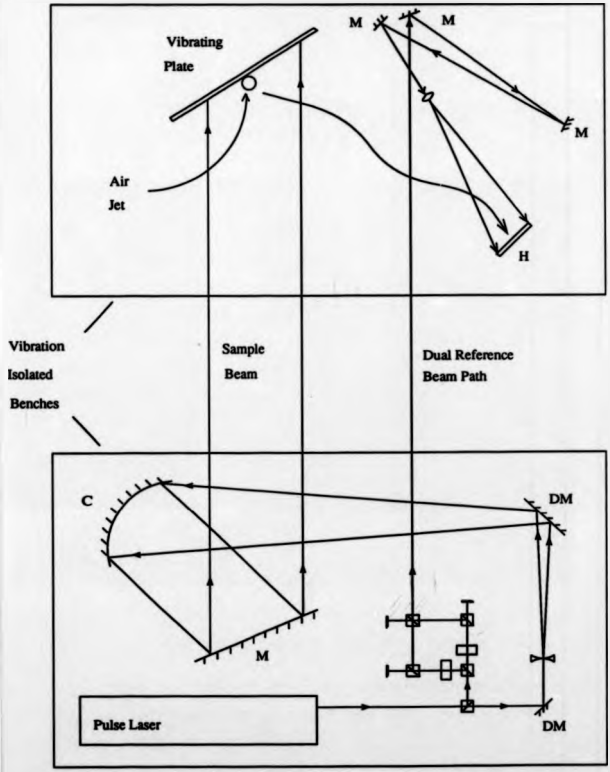
by dual reference beam holographic interferometry.

A similar system has been proposed by Lai [4]. In their arrangement a pair of acousto-optic modulators is used to switch the reference beams. One of the modulators is normally switched on and the other off. The state of both modulators is changed when a photodiode, placed in the hologram plane, detects the first pulse from the laser.

The full optical system used for vibration analysis is shown in figure 5.30. A collimated sample beam is produced to illuminate a vibrating sheet. Scattered light from the sheet combines with the reference beam at the holographic plate to form a transmission hologram. The two reference beams from the optical switching system are set to follow nearly identical paths. Their separation is controlled by rotation of the second 50/50 beamsplitter, and was set to approximately 200 microns. An arrangement of mirrors in the reference beam are used for path length matching before a short focal length positive lens spreads the beam over the holographic plate. All the optical components, and the object are mounted on vibration isolated benches.

The holographic interferograms were produced using a pulse separation of 250 microseconds. For this experiment there was no synchronisation of the laser pulse to the temporal behaviour of the vibration. Hence the measured displacements cannot be related to the amplitude at a given point. However the form of the vibration will be correctly represented in the fringe pattern. The main interest in this work was the evaluation of the holographic technique.

The holograms produced have been analysed using the single reconstruction beam method introduced in chapter 3. The image processing stages of pixel translation and filtering, formation of a wrapped phase map, and phase unwrapping have been performed. The results are shown in a sequence of images. Figure 5.31 shows the filtered and translated cosinusoidal fringe image (this image is the same as that used in chapter 1 to illustrate fringe tracking). The oval fringed area is the part of the vibrating sheet illuminated by the sample beam, with the air jet pipe approximately



KEY

DM - Dielectrically coated mirror M - Plane Mirror

H - Holographic Plate C - Collimator

Figure 5.30: Holographic Arrangement for Vibration Analysis

bisecting this area.

In this case a traverse calibration to a resolution of $\frac{1}{100}$th of a fringe was performed. Three images with phase steps of 120° were combined to form the wrapped phase map of figure 5.32. The image was analysed by the phase unwrapping algorithm using a tile size of 30 pixels. The results are shown in figures 5.33 as a normalised grey scale image, and figure 5.34 as a mesh plot.

Figure 5.31: Filtered Cosinusoidal Fringe Image of the Vibrating Sheet

In the same way as for the ESPI examples, an image showing the progression of the minimum weight spanning tree has been constructed. This is shown in figure 5.35 using the same notation as before.

The wrapped phase map shows a region of closely spaced fringes, immediately above the air jet pipe. These fringes are at the limit of the detectors spatial resolution. The period of the fringes in this region is approximately 5 pixels. This represents a

Figure 5.32: Wrapped Phase Map for the Vibrating Sheet

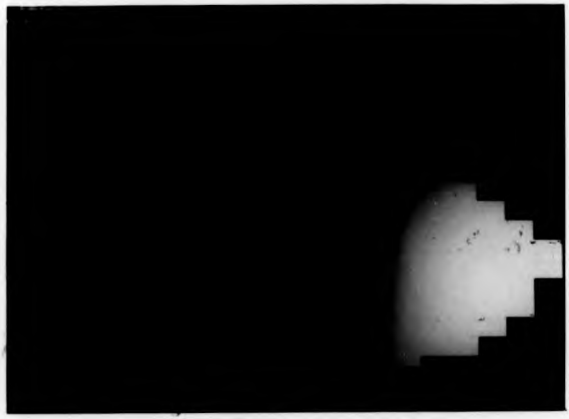


Figure 5.33: Unwrapped Phase Map for the Vibrating Sheet - Normalised Grey Scale Image

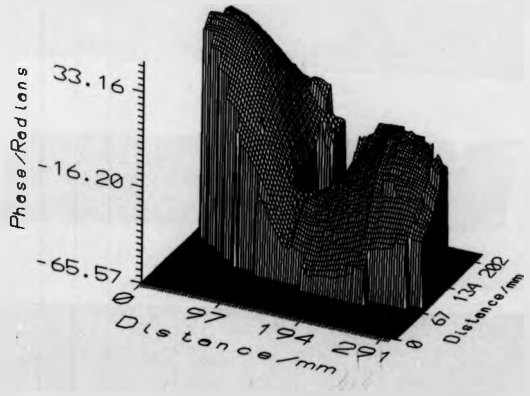


Figure 5.34: Unwrapped Phase Map for the Vibrating Sheet - Mesh Plot

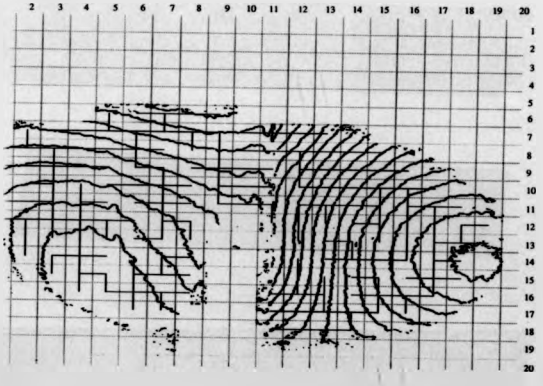


Figure 5.35: Progression of the Minimum Weight Spanning Tree in the Vibrating Sheet Image

more realistic resolution limit for a fringe than the Nyquist sampling theorem which states that two pixels are necessary to prevent aliasing of the data.

From the point of view of the phase unwrapping algorithm, the fringe field is effectively bisected by the air jet pipe. The majority of the data in each half is consistent. The dividing line exhibits fine fringes. It may be seen that the fringes are broadest at the top of the field, tile 11,8 in figure 5.35. Hence the tree connects the two halves here.

The unwrapped data shows the presence of the reference beam phase variation added to the deformation phase. In this case the precise slope of the ramp is unknown and a reference fringe image was not recorded. Therefore, only the first reference fringe compensation scheme can be applied, and an assumption regarding the actual deformation must be made. As the fringe field represents approximately one spatial period of the vibration a relatively unbiased estimate of the ramp may be obtained from the image. This approach may be used to assess the relative amplitude of a vibration across a surface when a translation of the surface has occurred at the same time.

The linear component of the phase distribution has been calculated using regression. The result of this analysis is shown in figures 5.36 and 5.37. Figure 5.36 shows the unwrapped map with the ramp subtracted (compare figure 5.34). The fringe order numbers are shown in figure 5.37 for the same cross section of the two unwrapped maps.

5.2 The Analysis of Phase Objects using Holographic Interferometry

A gas, or fluid, produces a change in phase of a light wave transmitted through it when the gas contains significant variations in density. This phenomena was quantified in

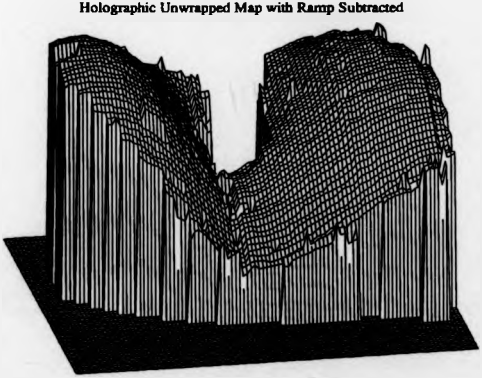


Figure 5.36: Mesh Plot of Ramp Corrected Unwrapped Phase Map

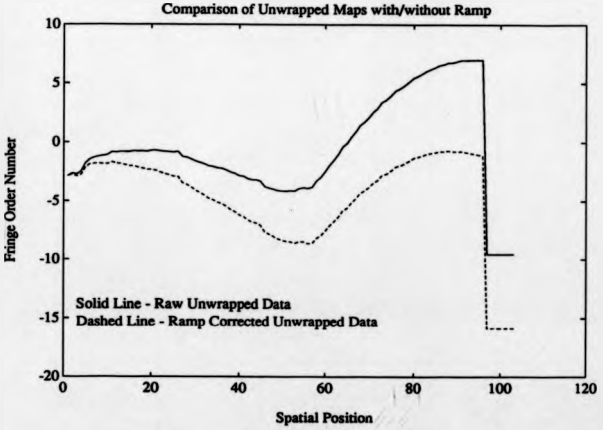


Figure 5.37: Cross Section of Original and Ramp Corrected Unwrapped Phase Maps

the Gladstone-Dale equation [5] :

$$n - 1 = K\rho, \tag{5.2}$$

where ρ is the density of the gas, K is the Galdstone-Dale constant, and n is the refractive index. Changes in refractive index then produce phase variations in a light wave transmitted through the gas. The phase changes of the interrogating wave are integrated along the line of sight and can be recorded holographically. It is assumed that the bending of the interrogating beam due to the refractive index field is either negligible or can be minimised. This is known as the refractionless limit.

The absolute state of a fluid may be assessed by producing the first holographic exposure without the flow of interest. The flow is then started and a second exposure taken. On reconstruction the fringe field maps the state of the flow against a uniform distribution of the gas at ambient conditions. Turbulent structures may also be examined by recording the two exposures, at two instants in time with the flow present. This is often termed a differential hologram, as the fringe field represents the unsteady component of the flow between the two exposures.

The fringe pattern may be described mathematically by considering the refractive index, n, along the ray path r for two holographic exposures. The phase change produced, $\Delta\phi$, at the image point (x, y) between the two states of the fluid is given by [6] :

$$\Delta\phi(x, y) = \frac{2\pi}{\lambda} \int [n_1(r) - n_2(r)]\, dr, \tag{5.3}$$

where 1 and 2 refer to the first and second exposures respectively.

Relating the phase distribution to a state variable of the gas is a complex problem. Phase changes are integrated along the optical line of sight through the flow. Therefore, in the general case, it is not possible to determine the density change at a particular point along the optical path. The full solution for a three dimensional density field can only be achieved by using at least three angularly separated views of the flow field. This requires the simultaneous recording of three holographic interfer-

ograms, and results in a complex optical system. The phase distribution from each viewing position must be solved. The integrated phase from each line of sight is then used as the input data, to form a tomographic solution for the density field [1].

Two simplifications are possible in certain cases which avoid the complexity of a multi-directional optical system, and the requirement for tomographic analysis.

Firstly, if the flow can be assumed to be two dimensional then a single view of the field is sufficient. The hologram is recorded with collimated illumination which is perpendicular to the principal direction of the flow. The refractive index of the fluid is then constant over the length of the beam path, L, and the integral in equation 5.3 can be simplified to :

$$\Delta\phi(x,y) = \frac{2\pi}{\lambda}\left[n_1 - n_2\right]L. \tag{5.4}$$

This may be re-expressed in terms of the fluid density by the Gladstone-Dale equation to give :

$$\Delta\phi(x,y) = \frac{2\pi K}{\lambda}\left[\rho_1 - \rho_2\right]L. \tag{5.5}$$

This statement relates the phase change seen at each point directly to the density change along the line of sight. Two dimensional flows have been analysed using this technique and the results used to 'fine tune' viscous flow solver codes [5, 7]. The analysis of the fringes was pre-dominantly done by hand using the stagnation point in the flow as a known zero phase change point.

This method has been successfully applied in the past, but current trends in the design of turbomachinery are towards three dimensional blade profiles and consequently three dimensional flows. Therefore, in practice, a two dimensional analysis of the flow will be applicable in fewer cases.

The second case, where tomographic analysis and multi-directional interferometers are not required is when an axis of symmetry is present in the flow to be analysed. This is principally found in the investigation of burner flames and also in plasma diagnostics. The refractive index field can be assumed to be radially symmetric and the Abel transform applied to convert the radial phase distribution into a refractive index distribution [8, p.267]. This analysis is applicable to a wide range of objects,

and implies that only a single view is required.

Any fluid which produces a variation in density can be investigated using this technique. Two cases are presented here. Firstly the examination of burner flames, where the density of the gas is mainly affected by temperature. An example is given of a simple laminar flame which has been solved to produce a continuous phase distribution. The single reconstruction beam approach has been used, to allow phase stepping of the fringe field. This data is shown to be at the resolution limit of a 512×512 pixel framestore. The methods which may be used to remove this resolution limit are discussed. The results from experiments to record multi-directional phase object information are also presented.

Secondly, three dimensional transonic air flows in a large scale transonic wind tunnel are examined. The visualisation of this data represents the current state of the art in this field. The possibilities for extracting quantitative data are then discussed.

5.2.1 Analysis of Burner Flames using Dual Reference Beam Holographic Interferometry

The analysis of burner flames can be approached in several ways depending on the type of flame examined, and the data which is required. For simple flames, such as a bunsen burner flame, the turbulence levels are low and an axisymmetric model can be assumed. In this case a single view through the object is sufficient. The solution for the interference phase along one line of sight can be used as input to the Abel transform to yield the density distribution of the gas.

In this work, the interest was in the optical recording systems, and analysis of the fringe data to produce a continuous phase distribution. The final stage of transforming the phase data into density, and subsequently temperature measurements is an area for future work.

A problem encountered at the reconstruction stage of a phase object is the depth over which the fringes can localise. If the fringes are localised over a large region of

space then it becomes difficult to image the fringes onto a photosensitive detector and produce clarity across the image. Large refractive index gradients are known to form fringes which localise close to the source of the disturbance. This condition can be used in some situations to determine the position of a sudden variation in the refractive index field [5]. However this effect is not desirable for fringe analysis methods to be applied across a whole field. The localisation of phase object fringes against a diffuse background was considered by Vest [8, p.284-288]. The results show that the depth of localisation is a function of the gradient of the field under examination. Therefore the localisation of phase objects can be expected to show variations across the field of view.

This effect has been reduced by placing a diffusing screen after the flame in the path of the sample beam. The phase change in the sample beam now occurs before the diffuser, and the fringes becomes localised close to the diffuser surface. This recording arrangement also causes the same fringe pattern to be scattered to all points on the holographic plate. In a conventional arrangement, with the diffuser behind the phase object, the fringe pattern will change with motion of the viewing point due to the different ray path which is observed at each point on the holographic plate. This is similar to the change in sensitivity vector which occurs with viewing point motion for an amplitude object. The problem is removed by scattering the light after the phase object. Therefore the single reconstruction beam method can be applied in this case without any restrictions due to changes in the fringe pattern with translation of the viewing point. The limitation of the method then lies in the quality of the imaging lens used.

In this work the diffusing screen has been replaced by a holographic optical element, or HOE. The use of a HOE has advantages in that the characteristics of the diffuser can be modified when the HOE is recorded. The optical system used to manufacture the HOE is shown in figure 5.38. A diffusing screen is illuminated by a collimated beam, and then imaged using a large aperture lens onto the hologram. By placing an aperture in the back focal plane of the imaging lens the angle of scatter

from the HOE can be limited. When the HOE is reconstructed the diffracted light is then contained within a limited scatter angle. This maximises the amount of light received by the hologram when the flame data is to be recorded and allows larger hologram to HOE distances. A second effect can be generated by placing a D.C. stop in the back focal plane of the imaging lens in figure 5.38. This prevents the directly transmitted light from the diffusing screen being recorded into the HOE. On reconstruction there is no directly transmitted light from the HOE to the hologram. This aids the localisation of the fringes onto the surface of the HOE and produces a higher fringe contrast.

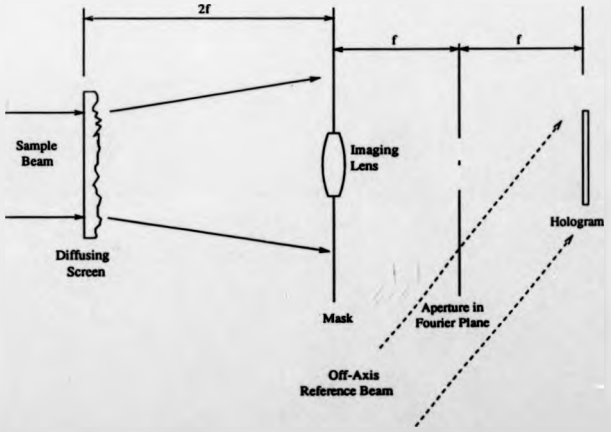


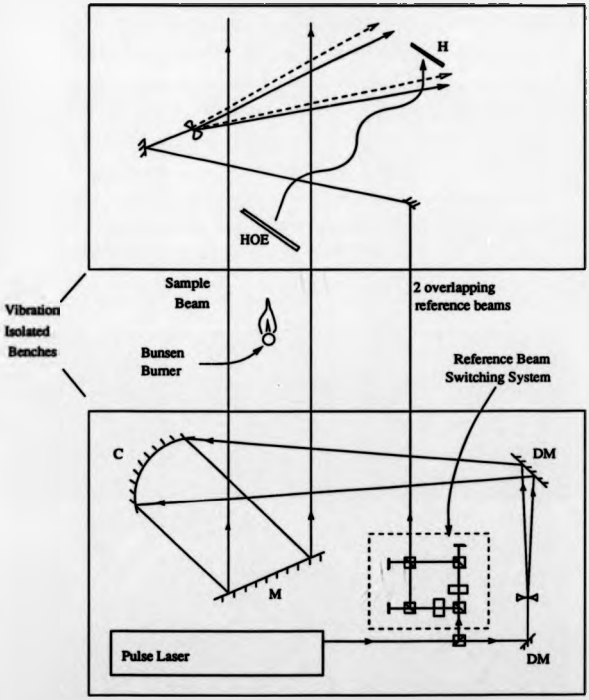
Figure 5.38: Optical Arrangement to Record a Holographic Optical Element from a Diffusing Screen

The main difficulty in using a HOE is to obtain sufficient diffraction efficiency. The

diffracted light from the HOE must be bright enough to form the sample beam of a second hologram. The diffraction efficiency of the amplitude transmission holograms described in section III is limited to 6.25% [9, p.56]. By developing the HOE to an increased density and then using a reversal bleach, the amplitude variations in the HOE are converted to local variations in the refractive index of the emulsion. This produces a maximum diffraction efficiency of 33.9%. Further increases can be obtained by forming a volume hologram through the depth of the emulsion, in which case a bleached hologram can produce a theoretical diffraction efficiency of 100%. The recipe for the bleach was obtained from Richmond Holographic Studios, a commercial organisation producing high quality, and large format display holograms.

A typical recording system to investigate small axisymmetric flames is shown in figure 5.39. A collimated sample beam is passed through the volume of the flame. This beam impinges on a holographic optical element (HOE) which acts as a diffusing screen with limited scatter angle. The holographic plate has been set parallel to the HOE in order to minimise the depth of the object which is observed. The optical switching system used for vibration analysis is shown to produce the two reference beams required. The two beams follow almost common paths to a negative lens and on to the holographic plate. The flame was located as far from the reference beam path as possible to avoid random phase changes being applied between the two exposures forming the interferogram. An absolute interferogram was formed, the first exposure taken with the flame lit, and the second after extinguishing the flame.

The hologram was reconstructed in a similar way to the metal disc data. Four phase stepped fringe fields were captured from equi-spaced points on the traverse. This data was analysed using the same sequence of operations as before, i.e. image translation in the framestore, filtering with a 3×3 pixel averaging operator, calculating the phase step between the fields, forming a wrapped phase map, and phase unwrapping. The filtered and translated cosinusoidal fringe image is shown in figure 5.40. The reference beam fringes can be seen on either side of the flame. The fringes at the edge of the flame show high spatial frequencies. This is caused by the



KEY

DM - Dielectrically coated mirror M - Plane Mirror
H - Holographic Plate C - Collimator

Figure 5.39: Dual Reference Holographic Recording System for Small Axisymmetric Flames

rapid increase in optical path length which occurs from the edge of the flame moving towards the centre. The wrapped phase map is shown in figure 5.41. The effect of the narrow fringes can be seen on the right hand side of the flame where small breaks are visible at the wrapover points in the fringes. The phase unwrapping algorithm has successfully analysed the data by solving the regions surrounding the flame first; the normalised grey scale image is shown in figure 5.42, and the mesh plot in figure 5.43. The two halves of the data are connected at the top of the flame where the fringes are of a slightly reduced spatial frequency. A small discontinuous region is present at the bottom of the flame where tiles have been joined to the tree from the inside of the flame. The extent of this region is limited by the consistency of the tile solutions with their neighbours.

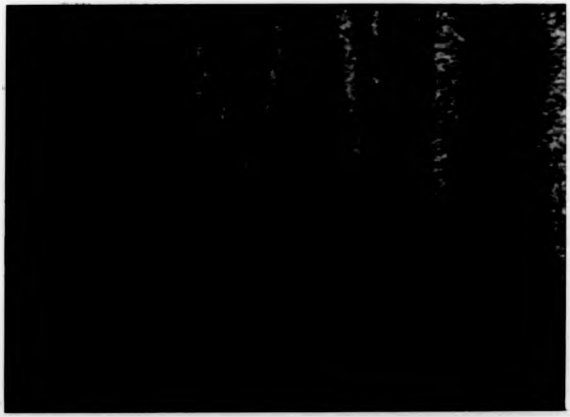


Figure 5.40: Filtered Cosinusoidal Fringe Image of the Laminar Flame

The fringe field consists of the central flame with ambient air forming the back-

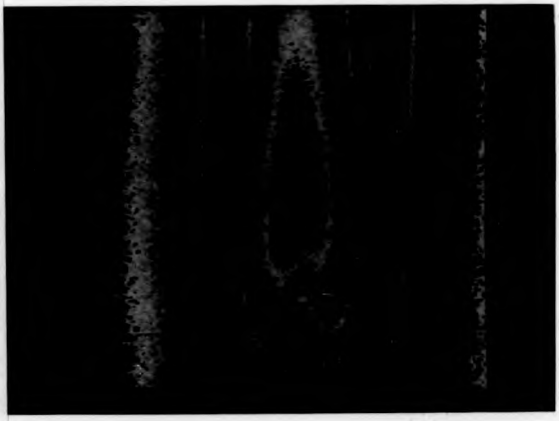


Figure 5.41: Wrapped Phase Map for the Laminar Flame

Figure 5.42: Unwrapped Phase Map for the Laminar Flame - Normalised Grey Scale Image

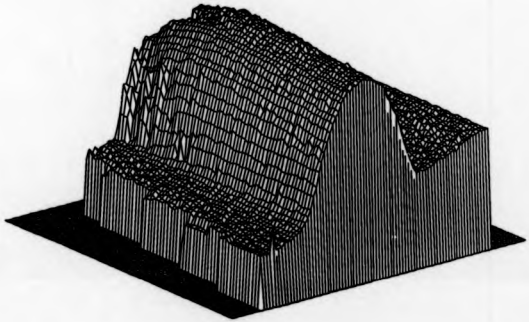


Figure 5.43: Unwrapped Phase Map for the Laminar Flame - Mesh Plot

ground on either side. The reference fringes have been assumed to be linear in order to determine the deformation phase. In this example the regions either side of the flame can be used to verify the applicability of this approach. After the linear component of the data has been subtracted the phase on either side of the flame should be equal. The resulting data is shown in figure 5.44 as a mesh plot.

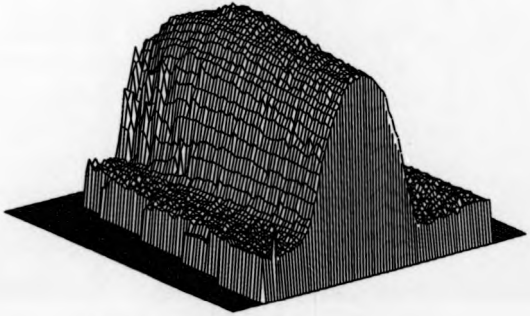


Figure 5.44: Mesh Plot of a Laminar Flame after Linear Ramp Compensation

This work demonstrates that phase objects which are resolvable by the currently available imaging devices can be successfully analysed using the single reconstruction beam phase stepping technique.

Methods for the analysis of more complex phase objects have been investigated. The first part of this work was to develop the optical systems which are required to

form multi-directional interferometric data. The aim is to produce three angularly separated viewing directions. Two approaches have been considered.

Firstly, the use of HOE's can be extended to form a multichannel HOE. A single incident beam on the HOE is diffracted into three components, each one formed from a diffusing screen. The phase object is placed in the region where the three diffracted beams overlap, see figure 5.45. There are two main disadvantages in this approach.

i) The fringe localisation will now form over a considerable depth. Therefore the distance of the interferograms and the viewing point from the mean localisation point must be large enough such that good fringe visibility is maintained across the whole flame.

ii) Secondly, the HOE itself must be large enough such that the flame can be positioned away from the HOE surface and that there is sufficient angular separation in the diffracted beams. This approach would require the specialist skills developed at Richmond Holographic Studios to handle large holographic plates (up to 1m square) and obtain uniform diffraction efficiency over the area.

If these difficulties can be overcome then phase stepped fringe data can be extracted from each hologram using the single reconstruction beam technique.

The alternative approach for producing multi-channel interferometric data is to form three image plane holographic interferograms with non-diffuse light. This requires large scale optics to manipulate three collimated beams through the region of interest. The use of specular interfering waves has the advantage of giving a higher spatial resolution in the images. In diffuse illumination work the spatial resolution is always limited by the speckle size. A second advantage is gained by using image plane holography. Refraction of the light waves due to large refractive index gradients in the flow are effectively removed by imaging the interfering waves into the image plane of the flame, which is made coincident with the holographic plate [10].

The single reconstruction beam method cannot be applied in this case as there is no depth between the hologram and the object. The application of the Fourier

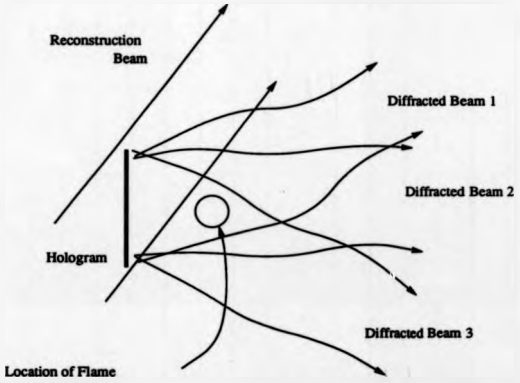


Figure 5.45: Use of a Multichannel HOE to form Multi-Directional Data of a Phase Object

transform method was considered. This approach was used by Bryanston-Cross [11]. Owing to the large number of fringes already present for this type of object, it is difficult to achieve a monotonic fringe distribution.

An alternative method to extract the phase data is to form two holograms for each viewing direction through the flame. One of the holograms is formed conventionally with the same reference beam for each exposure. The second hologram is formed with a phase shift inserted in the reference beam between the two exposures. The two fringe fields must then be digitised and correlated to occupy the same position in the framestore. The two image method introduced by Kerr et al [12] can then be used to form a wrapped phase map.

This idea has been demonstrated using two Pockels cells and a half wave plate, see figure 5.46. When both Pockels cells are off the wave undergoes a certain phase change through the half wave plate. With the Pockels cells on, the phase change is altered by a further π radians. Holograms have been recorded using this arrangement where a HOE was used to scatter the light rather than with specular wavefronts. Images from the two holograms are shown in figures 5.47 and 5.48. The flame data is seen to be of opposite phases between the two images. A phase shift other than π must be used in practice to form the wrapped phase map and resolve the directional ambiguity of the fringes.

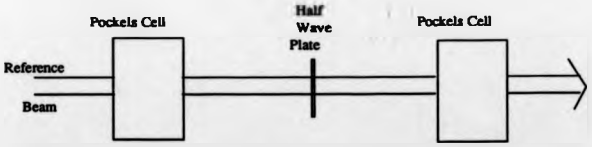


Figure 5.46: Optical System to Apply a 180 degree Phase Change to a Reference Beam

Assuming that the data can be recorded using one of the approaches described

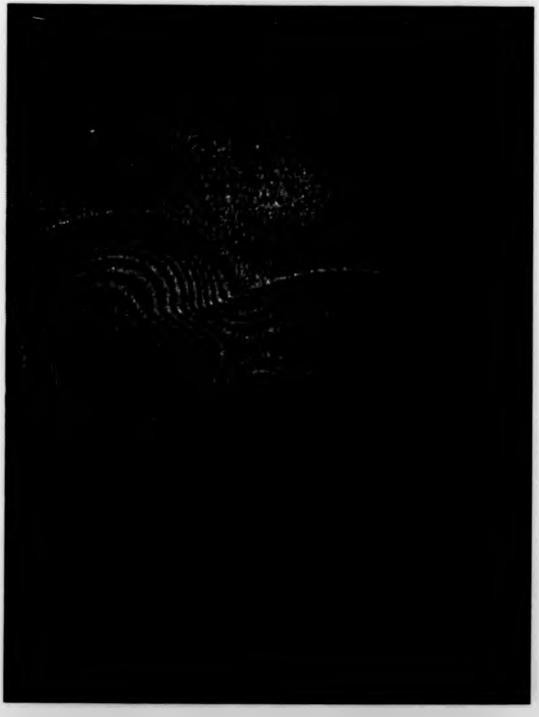


Figure 5.47: Conventional Double Exposure Interferogram of a Flame

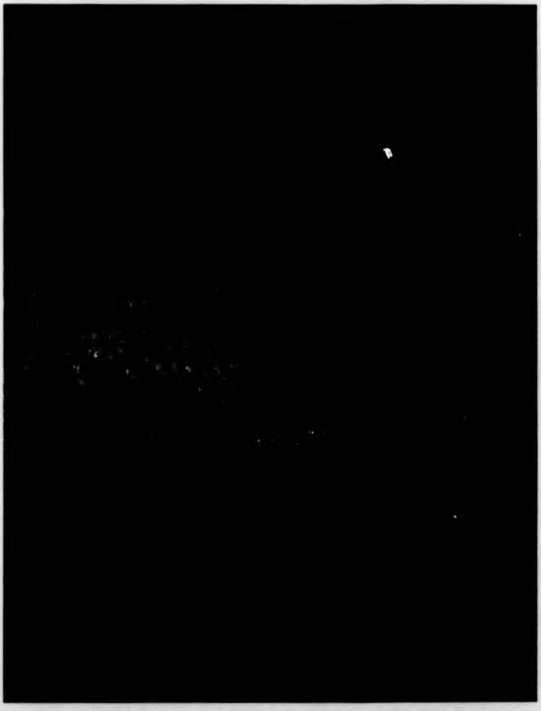


Figure 5.48: Double Exposure Interferogram of a Flame with Reference Beam Phase Modulation

above, a method to transform the fringes into digital form is then required. CCD devices and framestores are continually offering higher spatial resolutions, reports of 4096 x 4096 pixels systems are currently entering the marketplace. However the cost of these systems is high at this point in time.

The most promising option is to use a flat bed scanner. This technology has been developed primarily for desk top publishing applications for a system cost of $\approx$ $1k. The device can take an A4 sized photograph and digitise the data to a spatial resolution of 300 dots per inch with 256 grey levels. Some problems must be overcome for its successful application to fringe analysis : a compensation for the logarithmic photographic printing process is required to give true cosinusoidal fringes, and a means of registration is needed to overlay several phase stepped fringe patterns.

The ever increasing demand for image resolution ultimately leads to a requirement for more computing power to perform the analysis. The size of the data space for 4096 x 4096 pixel resolution is 64 times that of a 512 x 512 pixel image. With the advent of processors such as Intel's I860, however, there are possibilities that the analysis of high resolution fringe data could be achieved in an economically acceptable time.

5.2.2 Flow Visualisation in a Large Scale Transonic Wind Tunnel

The visualisation of transonic flows in large scale wind tunnels has been an on-going research project at Warwick University since 1985. The work was initially funded by the company concerned, the Aircraft Research Association at Bedford, and latterly as part of a DTI/LINK scheme.

The objective of the project was to produce a visualisation of the flow around a civil aircraft model in a large scale transonic wind tunnel. A cross-sectional view of the tunnel is shown in figure 5.49. The working section is 9 feet across by 8 feet in height. A typical model of half of a civil aircraft is shown with a turbine powered simulator (TPS) representing a gas turbine. The TPS is driven by compressed air

and produces a central core jet surrounded by a by-pass jet. The exit pressures of the two flows model the real case from a gas turbine. The specific flow features of interest were the interaction of the by-pass jet with the free stream flow across the wing surface, and the detection of secondary flows between the TPS and the wing. The flow is fully three dimensional and highly turbulent.

An understanding of the flow is required for the correct installation of a gas turbine onto an aircraft wing. A specific problem is to identify the characteristics which can cause the bypass flow from the TPS to become attached to the wing and thereby cause a loss of lift.

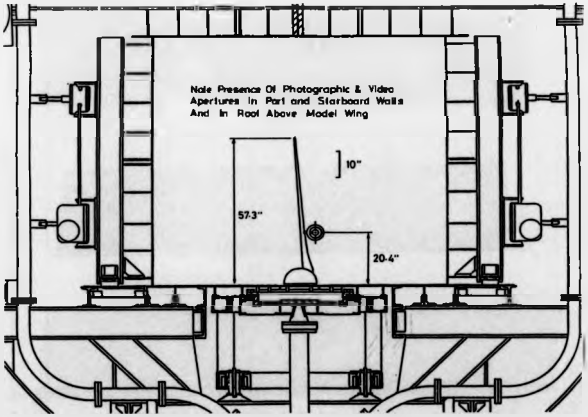


Figure 5.49: Schematic Drawing of the Transonic Wind Tunnel at the Aircraft Research Association

In 1987, prior to the start of this research, holograms of a static wing and engine were produced. The wing was rigidly mounted in a test cell, with flow being produced by a powered TPS unit. The hologram was recorded using the system in figure 5.50. A pulsed ruby laser was used, where the sample and reference beams were separated

by a prism. The reference beam was expanded onto a holographic film drive camera at the base of the wing. The sample beam was expanded from a position representing the side wall of the wind tunnel. Interferograms were recorded using a variety of pulse separations from the laser. A photograph from one of the reconstructed holograms produced with a pulse separation of 250 microseconds is shown in figure 5.51 [13].

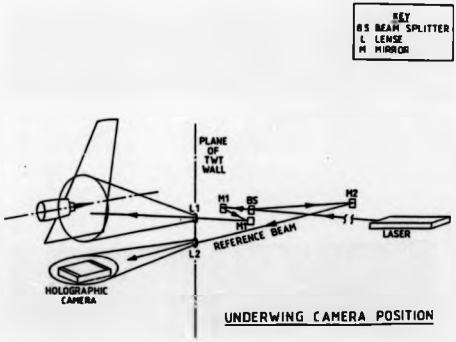


Figure 5.50: Holographic Recording System for Transonic Flow in a Test Cell

This hologram showed that a visualisation of the TPS flows could be made. The fringe pattern shows several distinctive features. Two visualisations of the jet flow have been made. The first is a two dimensional shadowgraph image formed by the projection of the sample beam through the jet. This is seen on the wing surface at the same level from the wing root as the TPS. Secondly, a three dimensional image of the flow is formed from the viewing point at the holographic film. This image is seen against the upper surface of the wing. The wing was painted white in order to enhance the amount of light scattered from its surface. The core flow is seen as a densely fringed structure emitted from the central cone of the TPS. The bypass

Figure 5.51: Reconstructed Image from an Interferogram of a Static Wing and Engine

flow is observed on both sides of the core with broader spaced fringes. The vibration patterns on the cowl of the TPS can be recognised. It is also possible to see large spherical waves centred on the TPS unit. These are acoustic waves caused by the high rotational speeds of the turbine.

This data showed the applicability of the holographic interferometry to visualise the flow. The application of a similar experiment to the transonic wind tunnel (TWT) was not straight forward. The major problem encountered was the operational costs of the wind tunnel, $\approx$ \$40k per day . This meant there was little opportunity to develop the holographic system during testing.

A second difficulty was to develop a suitable optical system around the wind tunnel. To prevent de-tuning of the laser during the test, the optical head of the laser was located in the control room of the tunnel. The beam path is approximately 19 feet from this point to the flow, and enters the wind tunnel through an armoured glass window. Beam manipulation optics were mounted on shelves along the wind tunnel side wall. Optical access into the working section for the sample and reference beams was obtained through 3 inch diameter holes.

Two systems were investigated for operation in the TWT. The first was based on the test cell work, see figure 5.52. A single expanded beam was directed towards the root of the wing. The edge of this beam passed through a neutral density filter and was incident on the holographic film camera mounted beneath the wind tunnel floor. The rest of the beam was scattered from the metal surface to the right of the film camera, up on to the model. Further scattering from the wing produces the sample wave recorded in the hologram. The sample beam was derived from a lower position than in the test cell work. This causes the shadowgraph and three dimensional images of the flow to be overlapped in the reconstructions. The clarity of the flow information is then reduced. In this system the reference beam is vulnerable to phase changes across the beam profile between the two pulses from the laser. This can be minimised by using short pulse separations.

The second system developed is shown in figure 5.53. The camera position is now in the side wall of the wind tunnel. The laser beam is manipulated behind the

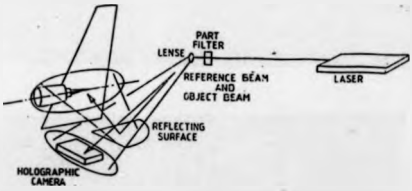


Figure 5.52: Holographic Recording System for Transonic External Flows - Wing Root Camera Position

wind tunnel side wall and a separate reference beam spread over the camera aperture. This produces a clean reference beam which is unaffected by the flow in the working section. The sample beam is expanded from the side wall directly onto the model. Scattered light from the wing is received through a window in the side wall, into the camera aperture. With this arrangement, the shadowgraph and three dimensional images of the flow are again overlaid, but the reference beam is undisturbed by the flow.

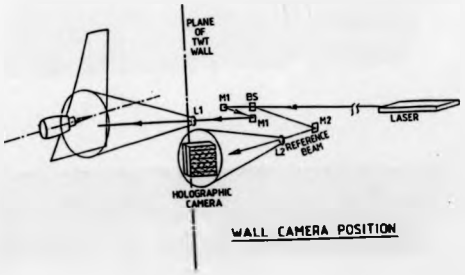


Figure 5.53: Holographic Recording System for Transonic External Flows - Side Wall Camera Position

Holograms were produced from both camera positions. The holograms obtained from the floor camera position showed fringes with reduced contrast. This was due to polarisation randomisation caused by scattering the sample beam off two surfaces. This problem could be overcome by using an appropriately orientated polarisation filter in front of the camera. Satisfactory images could not be produced from this data.

The camera position in the side wall produced some interesting results. The flow data was partially obscured by vibration of the wing surface in the background. This effect was minimised when short pulse separations were used from the laser. A photograph from one of the interferograms produced with a pulse separation of 20 microseconds is shown in figure 5.54. The wind tunnel was operating at a Mach number of 0.8. This image shows similar features as the result from the test cell. The core flow can be seen against the wing surface. On the cowl of the TPS unit a set of regularly spaced fringes in the bypass flow are observed. This fringe pattern was also observed from the other camera position, and in the test cell results. There is currently some debate as to the interpretation of these fringes. The acoustic waves can be clearly seen expanding from the back of the TPS unit. The angle of the waves is swept back in the direction of the incident freestream flow. Disturbances in the acoustic waves emphasise flow expansions from the edges of the TPS unit.

The clarity achieved in this result could be improved by further development of the optical system. The optics were assembled in the wind tunnel, and films produced from both cameras in less than half a day. The facility was operated for only $\approx$ 1hour. These results demonstrate the applicability of holographic interferometry to this type of industrial environment as a flow visualisation tool.

Further development of the holographic recording system is currently underway so that a better visualisation of the flow can be achieved. Holographic interferometry has been seen to offer a whole field visualisation of the unsteady flow field. Other visualisation methods cannot be applied in this case.

i) Oil film visualisation requires extensive model preparation time and only produces results at the model surface.

ii) Tracer particles can be added to the flow to perform a visualisation. Planar images of the particles are formed offering only two dimensional data.

iii) Schlieren and shadowgraph methods cannot be applied because of the location of the flow region of interest and the limited optical access to

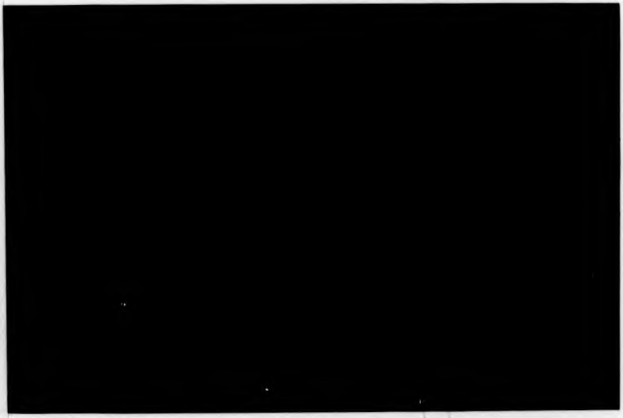


Figure 5.54: Reconstruction from an Interferogram of Transonic External Flows

this area.

The possibility of obtaining quantitative data from the flow using holographic interferometry is limited. The vibration of the rig as a whole currently prevents an absolute hologram of the flow being made. Therefore the holographic method can only be used to measure turbulence levels at this time.

The three-dimensionality and turbulence present in the flow imply that at least three angularly separated views of the field would be required to perform a full quantitative analysis. As the flow is imaged against a scattering surface, the fringe data will be localised over a large volume. Therefore it would be difficult to produce satisfactory images of the fringes. The data would also require converting into digital form to perform any analysis.

Therefore, alternative methods to obtain quantitative data from this flow have been investigated by other members of the group at Warwick University. Recent results have shown that particle image velocimetry (PIV) can be applied to transonic flows in large scale wind tunnels [14, 15, 16]. This method produces a direct quantitative measurement of the velocity distribution in a two dimensional plane using standard photographic techniques. The approach is simpler than holographic interferometry, as the two dimensional plane is imaged directly onto photographic film. There are other experimental difficulties for the implementation of PIV, as the flow must be seeded with sub-micron sized particles to provide a visualisation of the flow. The size of the particles is critical if the velocity data measured from the particle images is to be representative of the local flow velocity.

The quantitative analysis of three dimensional phase objects is a topic of continued research. The other main methods which can be used to obtain quantitative measurements from a flow are listed below [17, 5].

i) Laser doppler anemometers, or 2 spot anemometers can provide point by point velocity data. The operation time is typically 1 second per point; resulting in a time averaged velocity measurement.

ii) Laser induced fluorescence uses a planar light sheet from an ultra-violet laser to cause a gas component added to the flow to fluoresce. The intensity of the fluorescence is proportional to the velocity of the flow, thereby planar velocity fields can be produced.

iii) Particle image velocimetry also produces planar velocity measurements. Current research in this field is to extend the method to three dimensions, recording the particle image data holographically. Data reduction methods to track particle trajectories in three dimensions are also being explored.

Holographic interferometry can produce a fully three dimensional diagnostic for compressible flows. To obtain quantitative results further developments are required in the optical recording systems used. This may also necessitate the construction of specialist flow facilities with built in optical access.

Optical systems for dual reference beam holographic interferometry have been presented in this thesis. Thereby quasi-heterodyne analysis of the data may be possible. The data reduction of the fringe patterns requires high resolution imaging devices. This hardware is currently emerging into the marketplace. High speed computers are also needed to analyse the data. The algorithms to produce contiguous measurement data from the fringe pattern have been demonstrated on 512 × 512 pixel images. These methods are now becoming established and can provide automatic and reliable data extraction from digital fringe patterns.

The algorithms presented in this thesis are currently being applied by British Aerospace to solve interferograms of two dimensional flows [18]. This work shows a revival of interest from an industrial company in quantitative holographic analysis of transonic flows.

Bibliography

[1] D. Sweeney, C. M. Vest, "Reconstruction of the Three Dimensional Refractive Index Fields from Multi-Directional Interferometric Data",Applied Optics, volume 12, 1973.

[2] Tesa documentation on Inductive Axial Movement Gauge product number GT21. Tesa SA, Instruments et Systemes pour L'assurance de la Qualite, CH-1020, Renens, Switzerland.

[3] J. Davies, C. Buckberry, "Applications of a Fibre Optic TV Holography System to the Study of Large Automotive Structures", Society of Photo-optical and Instrumentation Engineers, volume 1162, San Diego 1989.

[4] G. Lai, T. Yatagai, "Dual-Reference Holographic Interferometry with a Double Pulsed Laser", Applied Optics, volume 27, number 18, p.3855-3858, 1988.

[5] P. J. Bryanston-Cross, "High Speed Flow Visualisation", Progress in Aerospace Science, volume 23, number 2, p.85-104, 1986.

[6] R. J. Parker, "Holographic Flow Visualisation", presented at a seminar 'Flow Visualisation - Its Application to Engineering Problems', the Institute of Mechanical Engineers, 21 January, 1987.

[7] P. J. Bryanston-Cross, "Holographic Flow Visualisation", Journal of Photographic Science, volume 37, p.8-13, 1989.

[8] C. M. Vest, "Holographic Interferometry", Wiley series in pure and applied optics, 1979.

[9] P. Hariharan, "Optical holography", Cambridge University Press, 1984.

[10] G. P. Montgomery, D. L. Reuss, "Effects of Refraction on Axisymmetric Flame Temperatures Measured By Holographic Interferometry", Applied Optics, volume 21, number 8, p.1373-1380, 1982.

[11] P. J. Bryanston-Cross, J. W. Gardner, "Holographic Visualisation of a Combustion Flame", Optics and Lasers in Engineering, volume 9, p.85-100, 1988.

[12] D. Kerr, F. Mendoza-Santoyo, J. R. Tyrer, "Extraction of Phase Data from Electronic Speckle Pattern Interferometric Fringes using a Single Phase Step Method : a Novel Approach", Journal of the Optical Society of America, volume 7, number 5, p.820-826, 1990.

[13] P. J. Bryanston, D. P. Towers, "A Holographic System for Aerodynamic Testing", IEE Holographic Systems, Components and Applications conference, Bath 1989 (published as IEE proceedings no. 399).

[14] P. J. Bryanston-Cross, A. Epstein, "The Application of Sub-Micron Particle Visualisation for PIV at Transonic And Supersonic Speeds", Progress in Aerospace Science, volume 27, p.237-265, 1990.

[15] P. J. Bryanston-Cross, S. P. Harasgama, C. E. Towers, T. R. Judge, D. P. Towers, S. T. Hopwood, "The Application of Particle Image Velocimetry in a Short Duration Transonic Annular Turbine Cascade", accepted for publication by ASME, July, 1991.

[16] C. E. Towers, P. J. Bryanston-Cross, "The Evolution of New Optical Diagnostics for Transonic and Supersonic Flows", seminar 'Flow Mapping using Laser Techniques', the Institute of Mechanical Engineers, 27 February, 1991.

[17] J. D. Trolinger, "Laser Applications in Flow Diagnostics", AGARDograph number 296, North Atlantic Treaty Organisation, 1988.

[18] Private Communication, S.Parker, K.Fisher, British Aerospace.

Chapter 6

Conclusions and Future Prospects

A novel, single reconstruction beam system, has been developed for the phase stepping analysis of dual reference beam holographic fringe data of diffuse objects. The technique offers practical benefits over conventional, dual reconstruction beam systems. The sensitivity to environmental variations is reduced, and the need for precision translation devices is eliminated.

Phase stepped fringe fields can be captured automatically by moving a camera on a linear traverse. Phase stepping algorithms using three or four images have been investigated.

The measurement accuracy obtainable in phase stepping analysis has been shown to be $\frac{1}{68}$th of a fringe. The accuracy of the measurements using the single reconstruction beam method and a three image phase stepping algorithm can produce an equal resolution in some cases, but never a higher resolution. The accuracy is reduced when the test object has significant height variations, and when the minimum fringe period is less than 16.25 pixels.

When four images are used to form a wrapped phase map, the phase stepping mechanism does not require calibration, thereby reducing the analysis time. The algorithm has been implemented such that the average phase step across the image is calculated. This parameter and three of the cosinusoidal fringe patterns are then

used to form a wrapped phase map. The accuracy of the phase measurements is reduced to $\frac{1}{40}$th of a fringe for this algorithm due to difficulties in obtaining an unbiased estimate for the phase step.

The single reconstruction beam system requires a feature in one of the phase stepped fringe images to be defined. A correlation can then be performed to overlay the positions of the fringe fields within the framestore. The successful operation of this procedure is dependent on placing a well defined object in the field of view.

Automatic algorithms for the analysis of interferometric data have been presented. Wrapped phase maps formed by the phase stepping and Fourier transform methods can both be analysed successfully over the whole field.

The methods were developed using holographic fringe images. The fundamental problem for the automatic analysis of wrapped phase data is the presence of bad data in the field, either at the pixel level, but especially over larger sections of the field. The presence of even a small number of bad data points can affect the phase calculated for the rest of the field when conventional unwrapping algorithms are applied.

It has been found that by making local measures of the fringe data, the influence of bad data can be restricted. Of particular importance is the local spatial frequency of the fringes. When the fringe spatial frequency approaches the Nyquist value the probability of obtaining erroneous data is increased. The phase data in the image is unwrapped from regions of low spatial frequency fringes. High spatial frequency data is only added to the solution at the ends of the unwrapping path when it cannot affect the rest of the field. This procedure forms a deterministic method to isolate regions of bad data from the rest of the fringe field.

Applications of the code to ESPI data were evaluated. A pre-processing algorithm has been demonstrated which automatically assesses the signal to noise ratio of the fringes such that the optimum degree of filtering is applied. The filtered ESPI patterns were found to be suitable for automatic analysis by the same phase unwrapping strategy. This work also showed the industrial potential of automatic fringe analysis systems for industrial inspection of components in the automotive and aerospace

industries.

There is no reason why the methods of automatic fringe pre-processing and phase unwrapping cannot be applied to other forms of interferometric data, for example, moire fringes.

The phase unwrapping strategy in its current form does not always produce optimum solutions. This is mainly because of the difficulty in automatically identifying the region of the fringe field to be analysed. The solution of the remaining problems will produce increased reliability in fully automated, whole field, fringe pattern analysis.

The applications of interferometric methods have been considered. Examples of static deflections using ESPI and holographic interferometry have been shown. The ESPI method offers rapid, near real time data capture, whilst holographic interferometry is limited by the wet processing of the photographic plates. Providing the data is resolvable by the imaging device, the fringes produced by both techniques are equally tractable using the automatic pre-processing and phase unwrapping strategies described.

The accuracy of the single reconstruction beam, holographic method was verified to the accuracy of an independent measurement probe ($\frac{1}{10}$th of a fringe) by monitoring the deflection applied to a metal disc. Further experiments are required to establish the method as a tool for nanotechnology scale measurements. Results from the analysis of simulated data show that measurements to 4 nanometers should be possible.

The single reconstruction beam holographic method has allowed a direct comparison between the phase stepping and Fourier transform methods of fringe analysis. The Fourier transform method was found to be sensitive to variations in fringe modulation and background intensity. Therefore, greater care is required in hologram recording for a Fourier transform analysis to be performed with equal accuracy as phase stepping. The results from the two processes showed a standard deviation error of $\frac{1}{16}$th

of a fringe.

Holographic methods are preferred over ESPI for the analysis of time varying events. With holography the data is digitised as a post processing operation and therefore the resolution of the image capture device can be selected accordingly. For ESPI, the data is only available at the resolution of the imaging hardware used to record the data.

The quantitative analysis of transient events using holographic interferometry has been achieved using a high speed optical switching system and a pulsed ruby laser. The optical switch allows dual reference beams to be used to record the data. The same reconstruction and phase stepping analysis methods have been applied to extract the phase measurements. The application of this system to vibration analysis was successfully demonstrated.

Applications of holographic interferometry to the quantitative analysis of phase objects have been considered. To obtain equal clarity of the fringe data across a phase object, it has been found necessary to image the flame onto a holographic optical element (HOE). The fringes are then localised at the surface of the HOE and can be imaged using conventional methods applied to amplitude objects. The solution of a simple burner flame has been presented where the data was resolvable in an image of 512×512 pixels.

The use of a HOE, to replace a diffusing screen in this case, has been found particularly beneficial. The characteristics of the diffusing screen were modified optically prior to recording the HOE. In this way customised optical elements can be manufactured for specific applications.

For the specific case of burner flames, it can be seen how to extend this approach to record multi-directional data, obtain solutions for the phase distributions, and hence from a tomographic solution to reconstruct the refractive index field. An alternative approach, offering higher spatial resolution data, is to form image plane interferograms with a specular sample wave. The fringe data can be digitised using a flat bed scanner.

Further development of these systems is expected to produce three dimensional measurements in complex burner flames.

The quantitative analysis of general transonic flows is not feasible at this time using the holographic methods described. The major problem to be overcome is the methods to image the fringe data against a diffuse background. Direct photographic methods, such as particle image velocimetry, offer a more direct solution to this problem at this time, producing two dimensional velocity data. However, in these applications, holographic interferometry is the only method to provide a whole field visualisation of unsteady events.

Chapter 7

Acknowledgements

The author wishes to express sincere thanks to the following people.

i) Dr. P.J.Bryanston-Cross for constant supervision and help during all stages of the work over the last three years.

ii) Mr. T.R.Judge, for his work on the fringe analysis software.

iii) Mrs. C.E.Towers, for putting up with 6 months hard labour and having the courage and necessary knowledge to proof read this thesis.

iv) The Science and Engineering Research Council for funding the majority of the postgraduate research.

v) Dr. Jerry Davies and Mr. Clive Buckberry of the Rover Group for supplying the ESPI data presented in this thesis and for many helpful discussions.

vi) The Aircraft Research Association Ltd., Bedford, where the holographic flow visualisation work was performed.

vii) Dr. Jurg Gass of the ETH, Zurich, for use of their holographic laboratory where the flame research was performed.

viii) Edwina Orr, and David Traynor of Richmond Holographic Studios for expert advice on chemical processing methods.

ix) Mrs. Sue Howard for proof reading the thesis.

x) Mr. D.C.Dyer for constructing the electronics used in the Laser Firing System of PBX Associates.

xi) Dr. S.T.Smith, for advice on measurement techniques.

xii) Mr. A.L.Hulme, for maintaining the computer system which we rely on.

xiii) The technicians of the Nanotechnology Group for constructing some of the equipment used in this research.

Appendix I

Characteristics of Laser Speckle

Speckle occurs when an optically rough surface is illuminated by coherent light. A surface is optically rough when its height variations are of the order of the wavelength of light. Considering the scattered light from such a surface in a plane at a distance z. Each point on the surface can be thought of as a source of spherical waves. At a point in the observation plane, light is received from all points on the surface. The phase of these scattered waves will vary randomly according to the local height variation at the surface. Consequently the resultant amplitude at the point in the observation plane will vary from zero to some maximum value as determined by the phase and amplitude of each component wave. When the observation point is moved, the phases of the scattered waves will vary, thereby giving a change in amplitude and the speckle effect.

Two characteristics of speckle are of interest in holography. Firstly, the intensity of a speckle. This has been shown by Goodman to follow negative exponential statistics. The probability, $P(I')$, of having an intensity greater than I' is given by [1, p.52-53]

$$P(I') = e^{-I'/\langle I \rangle}, \tag{I.1}$$

where $\langle I \rangle$ is the average intensity over many speckles. This equation shows that it is more likely to have black speckles in a field than bright speckles. This is seen in most fringe patterns : more dark speckles are seen in a bright fringe than bright speckles

in a dark fringe. This phenomena is useful in fringe tracking as it shows that there will be less noise present when a dark fringe is tracked.

Secondly, the characteristic size of the speckle is important. This parameter determines the spatial resolution of the measurements which can be achieved from rough surfaces. The characteristics of subjective speckle are considered here, i.e. the speckle size in the image plane of a lens as this is most relevant to the work which has been performed.

Consider the diagram in figure I.1. An optically rough surface is illuminated with laser light and imaged by a lens onto a photosensitive medium.

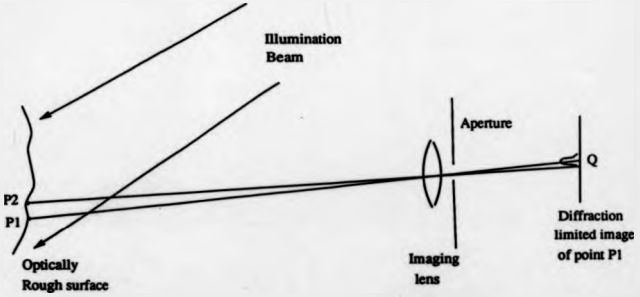


Figure I.1: Subjective Speckle formed in the Image Plane of a Lens

An object point P forms a diffraction pattern in the image plane of the lens, centred at Q. The intensity in this pattern follows a zero order Bessel function, i.e. a central peak surrounding by rings of decreasing intensity [2]. Light from a region around the point P is also imaged into the diffraction pattern centred at Q. These light waves each possess random phases as described above. The extent of the diffraction pattern is normally considered to the first zero of the Bessel function. Therefore light from an object point at P2 defines the radius of points whose scattered light can

affect the intensity at Q. The diffraction limited radius r_d in the image plane of a lens is defined as [2] :

$$r_d = \frac{1.22\lambda v}{d}, \tag{I.2}$$

where λ is the wavelength of light, v is the lens to image plane distance, and d is the diameter of the lens aperture. Therefore the size of the speckles, d_{sp} in the image plane can be taken as double this amount [1, p.56-57] :

$$d_{sp} \approx \frac{2.44\lambda v}{d}. \tag{I.3}$$

Therefore the speckle size in the image plane can be controlled by the aperture of the lens.

Bibliography

[1] R. Jones, C. Wykes, "Holographic and Speckle Interferometry", Cambridge University Press, Cambridge, 1983.

[2] M. Francon, "Diffraction", Pergamon Press, series on Coherence in Optics, 1966.

Appendix II

Solution of Quasi-Heterodyne Simultaneous Equations

The phase stepping methods which are considered in this thesis involve the use of three or four images with a fixed intermediary phase step. If the phase step is given by θ then the intensity at the same point in each field is given by :

$$i_0 = i_{dc} + i_m \cos(\phi + 0.\theta), \quad \text{(II.1)}$$

$$i_1 = i_{dc} + i_m \cos(\phi + 1.\theta), \quad \text{(II.2)}$$

$$i_2 = i_{dc} + i_m \cos(\phi + 2.\theta), \quad \text{(II.3)}$$

$$i_3 = i_{dc} + i_m \cos(\phi + 3.\theta), \quad \text{(II.4)}$$

where i_{dc} is the background intensity, i_m is the fringe modulation, and ϕ is the interference phase to be determined. The first equation can be re- expressed to give :

$$i_{dc} = i_0 - i_m \cos(\phi). \quad \text{(II.5)}$$

Substituting equation II.5 into the expressions for i_1, i_2, and i_3 and re-arranging each expression in terms of i_m produces :

$$i_m = \frac{i_1 - i_0}{\cos\phi\cos\theta - \sin\phi\sin\theta - \cos\phi}, \quad \text{(II.6)}$$

$$i_m = \frac{i_2 - i_0}{\cos\phi\cos 2\theta - \sin\phi\sin 2\theta - \cos\phi}, \tag{II.7}$$

$$i_m = \frac{i_3 - i_0}{\cos\phi\cos 3\theta - \sin\phi\sin 3\theta - \cos\phi}, \tag{II.8}$$

respectively.

The pair of equations II.6 and II.7 can be equated and re-expressed to give $\tan\phi$.

$$\tan\phi = \left(\frac{(i_2 - i_0)(\cos\theta - 1) - (i_1 - i_0)(\cos 2\theta - 1)}{\sin\theta(i_2 - i_0) - \sin 2\theta(i_1 - i_0)}\right). \tag{II.9}$$

This equation can be used with three images and a known phase step θ to calculate the interference phase ϕ.

The use of four images allows the phase step to be calculated also. The pair of equation II.6 and II.8 can be equated to give a second expression for $\tan\phi$. Whence

$$\tan\phi = \left(\frac{(i_3 - i_0)(\cos\theta - 1) - (i_1 - i_0)(\cos 3\theta - 1)}{\sin\theta(i_3 - i_0) - \sin 3\theta(i_1 - i_0)}\right). \tag{II.10}$$

Equating the expressions II.9 and II.10 and performing some algebraic manipulation gives a quadratic in θ :

$$\cos^2\theta\{2(i_2 - i_1)\} + \cos\theta\{i_0 + i_1 - i_2 - i_3\} + \{i_3 - i_2 + i_1 - i_0\} = 0. \tag{II.11}$$

The quadratic may be solved to give two solutions. Firstly a trivial solution :

$$\cos\theta = 1, \tag{II.12}$$

whence

$$\theta = 0. \tag{II.13}$$

This solution occurs due to the cancellation of a term $i_1 - i_0$ which implies a zero phase step. The second solution gives :

$$\theta = \cos^{-1}\left(\frac{i_0 - i_\theta + i_{2\theta} - i_{3\theta}}{2(i_\theta - i_{2\theta})}\right), \tag{II.14}$$

which allows the phase step to be calculated for each image point.

Appendix III

Processing of Holographic Plates

Holograms have been produced on Agfa-Gevaert 8E75HD silver halide holographic plates. This emulsion has a resolution of 5000 lines/mm. High resolution photographic emulsions are required in order to resolve the interference pattern between the reference and sample beams. In most of the experiments performed, the sample beam is normal to the holographic plate. Therefore the angle of the reference beam to the hologram normal, θ, determines the resolution required in the emulsion. The period of the fringes in the hologram plane, T, are given by [1, p.45-50]:

$$T = \frac{\lambda}{\sin\theta}, \tag{III.1}$$

where λ is the wavelength of light used to record the hologram. This corresponds to a spatial frequency of $\frac{\sin\theta}{\lambda}$. For $\theta = 30°$ and $\lambda = 694.3 \cdot 10^{-6}$mm the spatial frequency requirement is 720 lines/mm. This is a minimal figure and in practice to accommodate objects with large scattering angles, the resolution required approaches 2000 lines/mm.

The intensity ratio of the sample and reference beams in the hologram plane were set to $\approx$ 1 to 4, the reference beam being brighter. This ensures that the transmittance of the processed hologram is linearly related to the incident intensity, as described below. A typical transmittance versus exposure curve is shown in figure III.1, where exposure is defined as the product of the incident intensity and the exposure

time [1, p.46-50]. The exposure produced by the reference beam is set to produce an amplitude transmittance at the centre of the linear portion of the development curve. The ratio of the sample to reference beam intensities determines the amount of variation in the amplitude transmittance of the developed hologram around the centre of the linear section. The intensity of the sample beam must therefore be $< \frac{1}{3}$ of the reference beam in order to keep the variation on the linear part of the curve. If the exposure extends beyond the linear part of the development curve, the grating formed is no longer sinusoidal. This introduces extra diffracted waves from the hologram and higher noise levels in the virtual and real image waves [2, p.82-84].

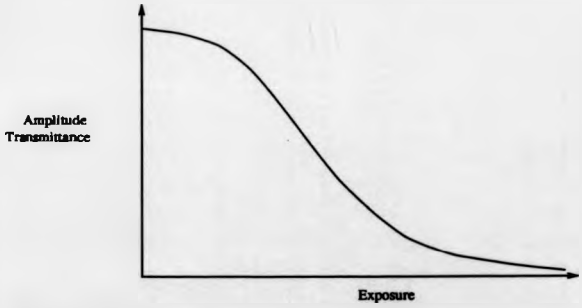


Figure III.1: Amplitude Transmittance vs. Exposure Curve for Holographic Emulsions

The holographic plates have been developed to give amplitude transmission holograms, i.e. the hologram effects the *amplitude* of the reconstruction beam which is *transmitted* through the hologram to give the diffracted waves. A typical development procedure is given in the table below III.

Process	Solution	Duration
Development	Agfa-Gevaert G128	≈ 2 minutes
	diluted 1 to 4 with water	
	or, Neofin Blue diluted	≈ 2 minutes
	1 to 2 with water	
Wash	Running Water	4 minutes
Fix	Kodak Unifix diluted	1 minute
	1 to 2 with water	
Wash	Running Water	4 minutes
Dry		

All the solutions are maintained at an approximate temperature of 20° Celsius. The developer converts the exposed grains of silver halide in the emulsion into black silver, producing a darkening of the plate. The fixing process removes the remaining, unexposed, silver halide to form a stable hologram.

The alternative to forming an amplitude hologram is to convert the developed silver into either a variation in refractive index or thickness of the emulsion. Upon reconstruction, the hologram effects the phase of the light rather than the amplitude. The diffracted waves are then several times brighter than for amplitude holograms, i.e. a higher diffraction efficiency is produced. In practice, it has been found that amplitude holograms are sufficiently bright for reconstruction with a CCD camera. The noise content in amplitude holograms is less than that in phase holograms. Therefore for holographic interferometry amplitude holograms are preferred.

Bibliography

[1] C. M. Vest, "Holographic Interferometry", Wiley series in pure and applied optics, 1979.

[2] P. Hariharan, "Optical holography", Cambridge University Press, 1984.

Appendix IV

Publications

i) D.P.Towers, T.R.Judge, P.J.Bryanston-Cross, "Vibration Measurements Using Dual Reference Beam Holography", SPIE proceedings 1084 'Stress and Vibration : Recent Developments', March, 1989.

ii) D.P.Towers, T.R.Judge, P.J.Bryanston-Cross, "A quasi heterodyne holographic technique and automatic algorithms for phase unwrapping", SPIE proceedings 1163 'Fringe Pattern Analysis', August, 1989.

iii) D.P.Towers, T.R.Judge, P.J.Bryanston-Cross, "Automatic Interferogram Analysis Techniques Applied to Quasi Heterodyne Holography and ESPI", invited paper to 'Optics and Lasers in Engineering', volume 14, p.239-281, 1991.

iv) D.P.Towers, T.R.Judge, P.J.Bryanston-Cross, "The Analysis of Holographic Fringe Data using the Dual Reference Approach", to be published in 'Optical Engineering', April 1991.

v) P.J.Bryanston, D.P.Towers, "A Holographic System for Aerodynamic Testing", IEE Holographic Systems, Components and Applications conference, Bath 1989 (published as IEE proceedings no. 399).

vi) D.P.Towers, P.J.Bryanston-Cross, T.R.Judge, "Development of a Pulse Laser system for the Production and Copying of Holographic Stereograms",

IEE Holographic Systems Components and Applications conference, Bath 1989 (published as IEE proceedings no. 399).

vii) D.P.Towers, P.J.Bryanston-Cross, C.E.Towers, "The Automatic Quantitative Analysis of Phase Stepped Interferograms", Society for Experimental Mechanics, conference proceedings on Hologram Interferometry and Speckle Metrology, November, 1990.

viii) P.J.Bryanston-Cross, S.P.Harasgama, C.E.Towers, T.R.Judge, D.P.Towers, S.T.Hopwood, "The Application of Particle Image Velocimetry in a Short Duration Transonic Annular Turbine Cascade", accepted for publication by ASME, July, 1991.

ix) C.E.Towers, D.P.Towers, T.R.Judge, P.J.Bryanston-Cross, "A Laser Light Sheet Investigation Into Transonic External Aerodynamics", SPIE conference proceedings, volume 1358, paper 192, High Speed Photography and Photonics, Churchill College, Cambridge, September, 1990.